



Online Communities of Creation as Collective Action. Access, Use, and Participation in a Digitalized Knowledge Economy

Nicolas Jullien

► To cite this version:

Nicolas Jullien. Online Communities of Creation as Collective Action. Access, Use, and Participation in a Digitalized Knowledge Economy. Economics and Finance. Laboratoire ICI, Information, Coordination, Incitation, 2013. tel-00960399

HAL Id: tel-00960399

<https://theses.hal.science/tel-00960399>

Submitted on 18 Mar 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Online Communities of Creation as Collective Action. Access, Use, and Participation in a Digitalized Knowledge Economy.

Nicolas Jullien

December 9, 2013

LUSI-iSchool, Télécom Bretagne, Brest. M@rsouin-ICI. Nicolas.Jullien@telecom-bretagne.eu

Document prepared in order to
present the
“Habilitation à Diriger des
Recherches” (HDR)
Université de Rennes 1.

Jury:

Patrick Llerena (Professor, University of Strasbourg), rapporteur

Dominique Torre (Professor, University of Nice Sophia Antipolis), rapporteur

Fabio Manenti (Associate Professor, University of Padua), rapporteur

Edmond Baranes (Professor, University of Montpellier)

Brian Butler (Associate Professor, University of Maryland)

Godefroy Dang Nguyen (Professor, Télécom Bretagne)

Thierry Pénard (Professeur, University of Rennes 1), directeur d’HDR

Abstract

This document presents the research I have undertaken over the last decade. It is both retrospective and prospective in the sense that, although it is obviously focused on my past activities, it also indicates ways for future research. The main topic of my overall research can be summarized as follows: I explore the development of online, open projects, or “communities of creation”, such as Free, Libre, Open Source Software (FLOSS), from an economics point of view. This means that in addition to renewing the answers to Olson’s question about the individual participation to collective action (1965), it questions also the why and how companies participate in this process, renewing Arrow’s dilemma (1962) on the incentives to produce innovation and the incentive to disseminate this innovation, and the way people organize themselves to transform participation into concrete pieces of knowledge, being software or encyclopedia articles.

Contents

Contents	3
List of Figures	5
List of Tables	6
1 Introduction.	8
1.1 Communities of Creation as an Object of Research.	9
1.2 Communities of Creation from an Industrial Point of View.	12
1.3 Individual Motivations and Recruitment.	14
1.4 The Market and Communities of Creation. A Work in Progress.	17
2 The Diffusion of FLOSS Products and Organization.	19
2.1 The FLOSS Competitive Advantage.	19
2.2 Firms' Strategies Regarding FLOSS.	22
2.2.1 Hardware.	22
2.2.2 Software and service.	23
2.2.2.1 Software platforms.	23
2.2.2.2 Packages.	23
2.2.2.3 Architects.	25
2.2.3 The dependence on users' skill.	26
2.3 The Role of Firms in FLOSS Communities.	26
2.3.1 Naïve users, the freeware strategy. The community as a commodity.	27
2.3.2 Business packages for VH users. The community as a core asset.	28

<i>CONTENTS</i>	4
2.3.3 Services to VH and KM users. The community as a complementary asset.	28
2.3.4 Conclusion.	29
3 Individual Involvement in Communities of Creation.	32
3.1 Beyond the Private-Collective Innovation Model?	35
3.1.1 The construction of the absorptive capacities.	36
3.1.2 The projects' environment.	37
3.1.3 Data and method.	38
3.1.4 Results.	40
3.2 Individual Profiles of Participation.	41
3.2.1 Surveying a community of creation. Data and method.	45
3.2.2 Results.	47
3.2.3 Discussion.	49
3.3 Evaluating Communities of Creation's Efficiency.	50
3.3.1 Model and data.	51
3.3.2 Results.	53
3.3.3 Discussion	58
4 Future Research: the Market and Communities of Creation: a Work in Progress.	61
4.1 Open-Source Demand.	61
4.1.1 (Information) Technology adoption.	61
4.1.2 Open-source ERP as a case study for open-source adoption.	63
4.1.3 Data collection.	65
4.2 FLOSS Business & Participation Strategies and Employee Management.	67
4.2.1 FLOSS Firms' strategies.	68
4.2.2 Employees' FLOSS strategy.	68
4.2.3 FLOSS firm's management strategies.	69
4.2.4 Model and data collection: the question of the measure.	70
4.3 Efficient Community, Good Contributor.	71
4.3.1 Measuring efficiency.	72
4.3.2 Looking for the "good" virtual team.	73

4.3.3	Participants' competences.	75
5	General Conclusion.	76
6	References.	78
6.1	Global References.	78
6.2	References Written by Nicolas Jullien.	104
6.2.1	Cited in the text.	104
6.2.2	Documents written by Nicolas Jullien not cited in the text.	106

List of Figures

1.1	Inputs, process and outcomes of online communities of creation.	11
2.1	The Compiere "ecosystem" (taken from: http://www.compiere.com/partners/)	25
2.2	Onion model based level of involvement in FLOSS communities targeted by firms regarding their business and the skill of their users.	30
3.1	Number of contributors versus Internet population.	53
3.2	Number of contributors versus population with a tertiary education.	54
3.3	Efficiency in recruitment of contributors.	56
3.4	Efficiency for the production of edits and articles (new articles and new FAs).	58
4.1	Open-source ERP adoption framework.	65

List of Tables

2.1	Firms' FLOSS strategy according to the skill of the users.	27
2.2	Firms' investment in the asset "community".	31
3.1	Research methodology and interview guide.	39
3.2	Repartition of the surveyed according to their level of contribution to Wikipedia.	46
3.3	Relationship between the self-evaluation of contribution and the time spent on Wikipedia. . .	47
3.4	Ordered probit with Heckman selection.	48
3.5	Return to scale for the recruitment of contributors.	55
3.6	Return to scale for the activities.	57

Acknowledgement.

It is quite a privilege to work in a school of engineering where research in social sciences is appreciated, since the foundation of the département d'économie et de sciences humaines by Godefroy Dang Nguyen, department where I did my PhD. The possibility to spend a sabbatical year in the USA, funded by Télécom Bretagne, facilitated the writing of this work. I am very proud to have been invited by Kevin Crowston at iSchool, Syracuse University, for this sabbatical, to have met Brian Butler there, and to have had him accepting to be part of my HDR jury.

This role of boundary spanner between a school of engineering and economics is quite challenging because of the risk of disconnecting from the academic community in one's field. Thanks to the support of researchers such as Jean-Michel Dalle, who was to first to be interested in FLOSS as an object of research in economics, Jean-Benoît Zimmermann, with whom I have collaborated for more than a decade, or Julien Pénin, with whom I started the reflexion on creative companies, bridges have been constructed and maintained. Without them, I would have dropped research long ago. But institutional networks helped a lot to meet these researchers: ESSID and ETIC European doctoral schools made me feel I belong to an economics research community. May those who helped me like organizers, be thanked! The fact that Patrick LLerana and Dominique Torre agreed to report on that work is a proof of the efficiency of such schools. I am very indebted to them. The inscription in University labs, such as ICI in Brest, and CREM in Brittany, as the M@rsouin research network have also helped to maintain the research links too, and I am proud to find my HDR director in this network. Thanks to him, Fabio Manenti accepted to report on this work, and Edmond Baranes, agreed to be part of the jury, may both of them be thanked for the time they spent on my work. Thanks also to Pr Pénard for having accepted to help me to get through this.

Chapter 1

Introduction.

Mobilizing hundreds (Linux) to thousands of contributors (Wikipedia), volunteer online open projects aiming at creating new knowledge, online “communities of creation”, as named by Rullani and Haefliger (2013) (i.e, communities aiming at producing a certain kind of knowledge, or “epistemic communities” (Cohendet et al., 2001), and doing it online), are viewed as central in the generation of new, innovative knowledge (Mahr and Lievens, 2012), in addition to providing new perspectives to the socio-economic question of participation in a collective action. The Olson paradox (1965) suggests that large groups are less able to promote their common interest than small ones because the individual incentives to contribute allegedly diminish with group size. However, many communities, in various contexts, have demonstrated their ability to develop selective incentives and institutions, allowing them to develop and protect their “commons” (Ostrom, 1990; Hess and Ostrom, 2006c). Moreover, these non-market projects dramatically changed the structure of some markets and industries, renewing Arrow’s dilemma (1962) on the incentives to produce innovation and the incentive to disseminate this innovation. Many scholars believe that this model of open, online collaborative knowledge production can and should be applied to other industries than software: Joly and Hervieu (2003) plead for a high degree of knowledge resource pooling in the field of genomics in Europe, not for renouncement of intellectual property, but by organizing a collective system of management of intellectual property. Hope (2008) defended the social advantages of an “open source” biotech industry. Dang Nguyen and Genthon (2006) called for “concentrating on an ambitious program of Free/ Libre Open Source Software (FLOSS)¹ production in the embedded systems

¹In this document, I will use FLOSS or Open Source Software indifferently. I define a “free, libre, open source software” as software distributed (made available), for free or not, with its source code and the right to modify the program and to redistribute

and domestic networks at the European level”. In both cases, the authors argue that this would reinforce the competitive position of European firms facing US multinationals and, by pooling basic technologies, avoid innovation clamping².

These two points are those I address in this work: what are the interactions between these non-market phenomena, or projects, and the market? (how can open business models be described, and what do they teach us about business models in general); what are the reasons why people participate in such projects and does firms’ participation impact these reasons and how open communities work? This will open the discussion about future research on what this organization teaches us about virtual organization, and also about cooperation between classical organizations and virtual ones. Before detailing these points, I will propose a framework to better understand the functioning of these communities, and to better explain the points I am studying here. Most of my case studies are looking at FLOSS and Wikipedia³ as, if Open Source initiatives are numerous, in various industries (Balka et al., 2009), they remain the most emblematic examples of such communities of creation. FLOSS is the movement which has impacted the most its industry and thus the case of reference in the open source IP model for innovation production. Wikipedia has become the most successful projects ever, with, for example, more than 4 million articles for the English version and one million visits per day.

1.1 Communities of Creation as an Object of Research.

What makes communities of creation specific, in comparison with, for instance, communities of practice, in the centrality of the creation of new knowledge goal in those epistemic communities (Amin and Roberts, 2008)⁴. Contrary to the communities of practice, the goal (the creation of a specific kind of knowledge) may precede the rules, or to paraphrase Wenger (1998, pp. 72-73), the ‘joint enterprise’ may precede the ‘mutual engagement’ (behavioral, interaction norms). Those communities are examples of “knowledge commons” Hess and Ostrom (2006b, p. 44), and their framework can be used to understand the production of such commons, distinguishing the characteristics of the community, or the input

these modifications. See Clément-Fontaine (2002, 2009) for a juridical analysis.

²On the debate about public intervention to support open-source initiatives, a debate I will not open here, see Horn (1999b); au Plan (2002); Comino et al. (2011).

³Olleros (2008) proposes a good introduction to the encyclopedia and how it has innovated in the production of encyclopedic knowledge.

⁴Those authors make a difference between communities of practice, and online communities. I must say that I am not very sure about this difference and would rather follow Cohendet et al. (2002) saying that communities of creation are a specific case of virtual epistemic communities.

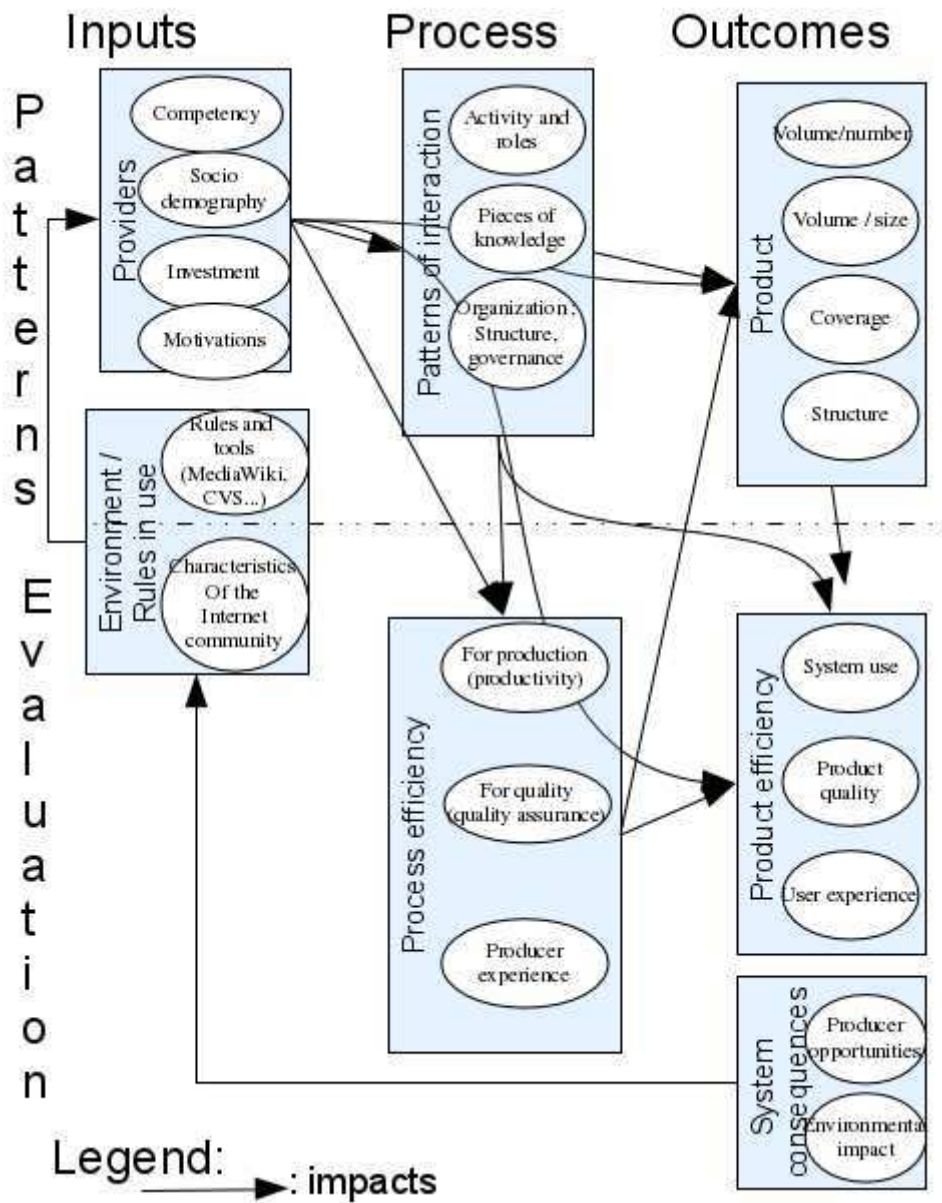
(“biophysical characteristics”, people or “attribute of the community”, “rules-in-use”), which constrain the way people interact (“the action arena”, or the process), leading to “outcomes” (the production). Of course, as mentioned by the authors quoted (*ibid*) the outcomes influence the inputs. The providers are given opportunities by their participation, leading them to potentially involve themselves more in the project; the users may also, by interacting with the system, become providers: for instance, Lih (2004) showed that Wikipedia articles cited by the press see the number of contributors increasing. One of the main differences between these online projects and the other common good productive communities is that the production outcomes (the pieces of software, the Wikipedia articles) are available to all, when the producers may have extra outcomes (and costs) to their involvement, as showed in Jullien et al. (2011), making the question of free riding even more accurate than in classical commons production groups.

Carillo and Okoli (2011), studying Wikipedia project, detailed Hess and Ostrom’s framework on the input and process part, even if they were less completely on the retroactions, and did not distinguish between the outcomes of the project and the specific outcomes for the participants, which is central to our analysis and to our work. For instance, Crowston et al. (2006), followed by Lee et al. (2009), propose indicators to analyze group production (they name “system creation”), and complete this model on two points. Relying on DeLone and McLean (1992, 2002, 2003), they proposed indicators to link the concrete outputs (in their case, open source software) to the user’s satisfaction. In their study, they also refer to Hackman (1987), to show the importance, as an output, of taking into account the producers (or contributors) feedback, and the process of development to have a global view of the outputs of such open online projects. They finally rely on Seddon (1997) to extend Delone and McLean’s model on the user side, with the concept of “perceived usefulness”, which echoes psycho-sociological studies on the adoption of systems by users, such as Technology acceptance model by Davis (1989) and its extensions (Venkatesh et al., 2003).

This leads me to a more global framework to study those communities (figure 1.1, page 11), where inputs are the providers like actors, the process, the action arena (action situations) and mainly the patterns of interaction, and the outputs, the outcomes, view from different viewpoints, users, but also producers (providers in Hess and Ostrom’s terminology), and which can be seen as an extension of the model proposed by Zhao and Bishop (2011, p. 720).

The patterns are the characteristics of the communities, and rely mainly on the analytic framework proposed by Hess and Ostrom: description of the inputs (who is participating, in which context), the

Figure 1.1: Inputs, process and outcomes of online communities of creation.



process, or action arena in their framework (how actors interact) and the products (outcomes). Based on the article cited before, I simply detailed in that scheme the evaluation criteria, which deal with efficiency, quality, user and producer experience. This is very helpful to organize the literature studying these communities (see, for instance the review of the literature on Wikipedia I proposed in Jullien, 2012), but also shows that a complete evaluation of such an organization is out of the scope of a single researcher and even of a single scientific field⁵.

From an industrial economics viewpoint, as said in the beginning of this document, the main questions are not in the understanding of the day-to-day life of those organization (roles, rules, process part) but in their relationship with the market: how the products impact the industry, how firms use them. And more generally, what are the relationship between the economic agents and those communities stressing here the question of the individual investment in those communities. In a word, the emergence of those communities are re-questioning Olson's question about the individual participation to collective action (1965), and Arrow's dilemma (1962) on the incentives to produce innovation and the incentive to disseminate this innovation. Because of this temporal and thematic continuity with my PhD work, because of this impact, I will, first, propose an analysis of the links between the FLOSS organization and the firms.

1.2 Communities of Creation from an Industrial Point of View.

As said before, FLOSS remains the movement which has impacted the most its industry and thus the case of reference in the open source IP model for innovation production (Jullien and Zimmermann, 2002). the economic success of FLOSS, and the difficulties to imitate it in other industries, may be explained

⁵For the sole domain of Wikipedia, there were 7029 articles in the science direct base on October 20th, 2011, notably because this encyclopedia is used as a test base in information and language processing systems (See, for instance, the researches conducted at University of Amsterdam, <http://ilps.science.uva.nl/search/node/Wikipedia>) and information retrieval tasks (Buriol et al., 2006). Still using Wikipedia, and without claiming to be exhaustive, here are some questions raised by these socio-technical projects (Bryant et al., 2005; Benker and Nissenbaum, 2006), where the tools used and the rules mediate and shape user activity around open collaborative writing, which can be seen as a community of practice (Hara et al., 2010), or even as an aggregation of multiple communities of practice (see, for instance, the analysis of the use of Wikipedia by sport fans by Ferriter, 2009). Regarding its functioning, Okoli (2009); Park (2011); Okoli et al. (2012) may have proposed the most recent review of the literature, which can be split into three main themes: motivations to contribute (Nov, 2007), and link between these motivations and the quality of the contribution (Glott et al., 2010b); editorial process or internal organization (Besten and Dalle, 2008; Brandes and Lerner, 2008; Fréard et al., 2010; Kittur et al., 2007b,a; Ortega and Gonzalez Barahona, 2007) and its impact on quality (Viégas et al., 2007a,b; Okoli and Oh, 2007; Stvilia et al., 2008; Carillo and Okoli, 2011), with a majority of articles in Information System (IS), Computer Mediated Communication and Computer Supported Cooperative Work; quality and reliability of the production, with more communication and library science (Denning et al., 2005; Magnus, 2006; Svoboda, 2006; Gorman, 2007; Waters, 2007; Fallis, 2008; Dede, 2008; Fiedler, 2008; Eijkman, 2008; Rector, 2008; Santana and Wood, 2009; West and Williamson, 2009; Royal and Kapila, 2009; liang Chen, 2010) and teaching orientation (Callis et al., 2009; Haigh, 2011), with more critical studies before 2007, even if Giles (2005) is the first publication which proposed a comparison of both Wikipedia and a classical encyclopedia, in favor of the first.

by the fact that software is quite specific in terms of good produced and organizational model (Mateos-Garcia and Steinmueller, 2003), and thus that the other industries failed to understand it well enough to adapt it to their situation. In a recent analysis of the transferability of this model, Pénin (2011), explained that it is a very new model of innovation, the "open-source innovation model", different from Chesbrough's open innovation model (2003; 2006), "both more open and more interactive than open innovation", as the dynamic of interaction is more based on the bazaar, or on the Cohen et al.'s garbage can model (1972), according to Lacolley et al. (2007); Li et al. (2008), and not on a strategic exchange of certain part of intellectual property goods. In a word, FLOSS is said to propose an original solution to the segmentation of knowledge production and is apparently more efficient than traditional Intellectual Property (IP) systems in cases where knowledge is modular and cumulative, as it facilitates and decreases the cost of producing new knowledge (Bessen, 2005). This is because "entry competition and innovation may be easier if a competitor needs only to produce a single better component, which can then hook up the market range of complementary components, than if each innovator must develop an entire system" (Farrell, 1989).

The FLOSS paradoxical situation, where commercial business relies on the existence and durability of non-market activities, questions industrial economics. This has clearly something to do with "cooperation" questions (Brandenburger and Nalebuff, 1996). As in any cooperative agreement devoted to technology or knowledge development, agents put assets together in a "pre-competitive" phase and share the products of their efforts before coming back to competition (Crémer et al., 1990; Bhattacharya and Guriev, 2006). On the contrary, a FLOSS project is an open game in which the list of players is not bounded ex-ante by a cooperative agreement and whose product is a public good that cannot be privately appropriated by the players. This would rather correspond to the formation of an open consortium for the production of a standard⁶, as studied by Coris (2006) and Simcoe (2006), in a Callon's dynamic innovation (1994; 1999) context (Pénin, 2011). Relying on the dynamic framework of innovation built by Callon (ibid), this author explained that in the beginning phase, open source innovation model can be considered, as "a platform, a springboard which business firms can tap into to improve their innovative ability" and where "firms and communities of individuals collaborate to feed an open pool of knowledge, which can then serve as a basis for corporate competition", echoing Jullien and Zimmermann's view (2006) of these

⁶What we mean is that a player offers a standard by developing a software the other players can adopt and contribute to the development. This "unilateral" adoption is usually called 'bandwagon' in the literature on standards (see, for instance Farrell and Saloner, 1985).

open source platform as a Romer-like industrial public good (as for other situations where actors try to capture innovation, see Cohen and Levinthal, 1989). His conclusion is that this model seems particularly adapted to the creative industries, such as the music or the game industry, studied by Bach et al. (2010).

Following what I did in my PhD, I propose in this document, and specifically in chapter 2, first to understand how and why FLOSS products and organization succeed in the context of Internet diffusion, before coming back to the reasons for companies to get involved in such projects, as, still quoting Iansiti and Richards (2006), in this cluster, "significant investments have been made in projects that will serve as complementary assets to drive revenues to vendors' core businesses"⁷. I will defend the idea that a key element to understand Firm's FLOSS business model is the skill of the users, before showing that this framework helps to understand the diffusion of FLOSS within the market, and proposing an industrial economics analysis of the involvement of the firms into FLOSS production by their business model, based on Teece's theory of the dynamic capabilities.

Quite logically, if the involvement of the market into the production of open, online communities is that dependent on VH users, the stability of such communities depends on their capacity to attract such users, and to make them participate. That is the reason why, in the second part of my research, I tried to better understand the motivations, the characteristics and the behavior (activity and roles) of these open online production projects participants.

1.3 Individual Motivations and Recruitment.

In classic knowledge or software production systems, there are either financial flows from the user to the producers, or collective, public support of the producers (as in open science) to do so⁸. In this regard as well, these open communities seem innovative, providing a new perspective to the socio-economics question of the participation in a collective action, for the creation of knowledge as public good, better known as Olson's paradox (1965): without direct monetary retribution, although it is easier to free-ride than in classical public goods production (Hess and Ostrom, 2006a), there are enough contributors to make the project work.

⁷In that respect, Lakhani and Wolf (2005), analyzing the results of an investigation of 684 software developers involved in 287 FLOSS projects, show that "a majority of (their) respondents are skilled and experienced professionals working in IT-related jobs, with approximately 40 percent being paid to participate in the FLOSS project".

⁸Foray and Cassier (2001b,a) propose a synthesis of the discussion of the ins and outs of the economics of knowledge creation and insist on the need for the creation of incentives for the producers of knowledge to produce and to diffuse.

But there is a recurring concern about the difficulty of recruiting and retaining new contributors, noted and studied by Von Krogh et al. (2003) in the case of open source software community, or by Halfaker et al. (2011) for Wikipedia. This process of joining is constantly needed, first because there is a decrease in involvement with the number of years people are active in the community (Ortega and Izquierdo-Cortazar, 2009 for open online communities, Borzillo et al., 2011 for intra-organization ones) and secondly because the production of knowledge needs a constant renewal of ideas (Arora and Gambardella, 1994). Specific programs, targeting newcomers, have even been designed to mentor those newcomers (see Musicant et al., 2011, for Wikipedia), and a lot of studies have been conducted to better understand the motivations for contributing (von Krogh et al., 2012 proposed a review of these motivations and of the literature studying them), as the different steps to join the core (Von Krogh et al., 2003; Hertel et al., 2003, and more recently Crowston, 2011).

Person's role evolves over time (Von Krogh et al., 2003; Jensen and Scacchi, 2007) in those projects, as do the reasons why s/he participates (Von Krogh et al., 2003). For instance, Shah (2006) showed that long-term participants enjoyed programming and interacting with the rest of the community (i.e., labeled as "hobbyists"), whereas short-term participants were typically driven by an immediate need for software (i.e., use value). Communities of creation have been said to organize the collaboration between actors of divergent interests (O'Mahony and Bechky, 2008), creating some "coat-tailing systems" to integrate heterogeneity in terms of contributions and goal (Hemetsberger and Reinhardt, 2009). More generally it echoes the results of the critical mass theory regarding the construction of collective action (Oliver et al., 1985; Marwell and Oliver, 1993), and the theory analyzing the construction of the (knowledge) commons (Ostrom, 1990; Hess and Ostrom, 2006b): these projects are made possible by the aggregation of various motivations and level of involvement⁹.

But, according to von Hippel and von Krogh (2003), beyond all these motivations, the core of the incentive framework is the "private collective" innovation model or the "user-as-innovator principle" (Lakhani and von Hippel, 2003; von Hippel and von Krogh, 2003): as users directly benefit from the piece of innovation they produce, they have incentive to produce it, and as they can expect add-on, feedback or cumulative innovation on their own proposition, they have incentive to freely share it. von Hippel and von Krogh's view can be interpreted as the fact that the only industries where FLOSS-like organization may flourish are those where producers and users are the same. To put it in a nutshell, even if many

⁹For a formal model of this phenomenon, see Rahman (2008)

reasons to participate exist, evolve with time and participation, these incitations would come only in addition to the main incitations which is the use. The direct consequence is that, regarding the building of open online communities, this would indicate that only communities where producers are users can succeed. Is it true? Is this motivation present since the beginning or is it more an experience good? Do the projects, even the successful ones, provide the same level of incitations, are equally efficient in attracting the contributors? Those are the questions I recently discussed and I will present in the third chapter.

To test the producer-user theory, I studied domains where FLOSS organization should be, *a priori*, of maximum efficiency, because the production of knowledge is modular and incremental, but where the users and the producers are disjointed. This exploratory work has been done via a qualitative approach (Jullien and Roudaut, 2012). The conclusion of this work is that it seems that the producers have to be users. This questions the way people get involved in those communities, and if this investment is made by a certain kind of user. There is a growing concern of community managers about the difficulty of recruiting and retaining new editors, in Wikipedia (Ortega, 2009; Halfaker et al., 2011), leading to the development of specific recruitment programs (Musicant et al., 2011), but already stressed and studied by Von Krogh et al. (2003) in the case of open source software community joining. The survey of the users of the French Wikipedia project I conducted in 2011, and the analysis of the more than 13,000 responses obtained via econometric tools (Probit and Heckman models) showed that the intensity of the first contribution, the time period between the discovery of Wikipedia and this contribution teaches a lot about the future level of involvement in the community (Dejean and Jullien, 2012). The will from certain communities to improve their efficiency in recruiting new contributors somehow means that it is possible to evaluate and compare the projects on this aspect. I proposed a methodology to evaluate and compare the efficiency with which the different language Wikipedias turn their readers (inputs) into contributors (outputs) and the contributors (inputs) into articles (outputs), based on Data Envelopment Analysis (Crowston et al., 2013b).

All these points will be presented in chapter 3. The question of efficiency is important, not only for the managers of such communities, but also for the institutions which want to get involved: they cannot invest in stalled projects, or projects where the return is decreasing. This leads to the third point I want to discuss in this document: what research is to be conducted to better understand the link between institutions and online communities, participation in collective production and involvement in a company, collective work and job.

1.4 The Market and Communities of Creation. A Work in Progress.

Since I began my journey in the communities of creation's world, a lot of work has been done to better understand how they work, the reasons why people participate instead of free-ride, and the relationship between these communities and other systems of exchange, such as the market. But there is still work to do, especially on the managerial impact of such a relationship. These communities of creation can be viewed as new ways of organizing the production, but also as a new way of producing and delivering goods to the market. This consideration is at least threefold, three directions in which I would engage research in the near future.

A first question, surprisingly much neglected in the literature, is the reason why clients buy open-source-based offers. Actually, the adoption of open-source products renews another industrial economics classical question, about the reasons and the path of diffusion of a technology. As shown by Dedrick and West (2004), the overall adoption of an OSS product in the market can be explained by economic diffusion theory (e.g., due to increasing returns due to investments made in hiring and training skilled IT workers, perception that the open-source solution will be the standard), while each individual adoption can be predicted by diffusion-of-innovation theory ("compatibility with current technologies and skills, organizational resources and tasks, and the availability of external technological resources", p. 9). Regarding the adoption of the technology, in a Roger's perspective, the fact that an OSS package can be tested before adoption (i.e., trialability) seems to be a strong benefit for these offers. In addition, and already said here, the process of production proposes a new solution between norm and standard (Jullien and Zimmermann, 2006b; Simcoe, 2006), as the standard is continuously discussed by several producers. Finally, in several markets where open-source products meet success, there are one or more strong proprietary solution standards: data bases (Oracle, which bought an open-source competitor, MySQL), operating system (Windows and Unix vs. Linux), and ERP (SAP vs. Open Bravo, ERP5, Compiere). But there are few evaluations of the reasons why companies choose open source solutions, and I propose here a theoretical model to test the reasons for adoption.

Secondly, regarding the producers, the industrial framework for the FLOSS industry (Jullien and Zimmermann, 2011a,b, b) has to be extended: Does the evolution of the market (big data services, cloud computing, more generally the software as a service trend) challenge this model, and is it validated by the behavior of the firms, regarding the participation in open online projects? Looking at my work on the

subject and at the recent literature, the management of the relation between an open-source firm and the related communities is of crucial importance. Ågerfalk and Fitzgerald (2007) have observed that to preserve the co-existence and co-operation of two types of organization that are based on remote albeit not contradictory rationales, firms must, in a nutshell, “not seek to dominate and control process”, “provide professional management and business expertise” and “help establish an open and trusted ecosystem”. They view such interaction as rather more osmotic than parasitic, as the firm’s resources reinforce communities sustainability. All this questions the impact of communities on firms’ organization and management. How should firms organize themselves to capture the feedback from the communities? The main questions here are as follows: in a labor economics perspective, why do companies let their employees participate in open projects and how do they manage their involvement? Is this involvement guided by strategic purpose only (as the employees represent the investment of the firm in to project), or are other points at stakes, such as the training of the employees, the negotiation of of some compensations (perks) to attract high profiles? On the other hands, are open-source participants using their involvement to signal their high profile?

But the origin of the open-source rationale remains that of developer-users pooling their development efforts for their own needs, aiming at better access to efficient tools for everyone. Whether this volunteer collaboration is initiated by individuals or by a firm, a key condition of success is users-developers adhesion. So, and following up the work I started on the measure of the efficiency of communities, it is of growing importance to better understand how these groups attract and retain contributors, and also turn their contribution into “tangible” production (articles, pieces of code...): what the characteristics of the teams producing these pieces of knowledge are, what the roles participants take in those communities are, and if these cultural, intellectual, social benefits developed by the participants are transfered onto the job market. I will detail how future research may be conducted in those aspects in the future in chapter 4.

Chapter 2

The Diffusion of FLOSS Products and Organization.

Industrial economics (Shepherd, 1990) explains that an industry is characterized by the basic conditions of each kind of activities: characteristics of the products, of the users, hence of the demand, but also of the juridicial environment (intellectual property protection, for instance). These basic conditions draw the main aspects of the market structure (source of added value, competitive advantages, sources of barriers to entry) and the nature of the competition (firms behavior, in terms of price, position, etc.) The efficiency of the firms (their performance) would appear to depend on a good adequacy of their strategy (behavior, organization) to the market structure, and to their capacity to reshape this market structure, increasing, for instance, the barriers to entry (Tirole, 1989b).

In this part, I propose to study the structure of the computer industry and to explain how this structure has evolved to create conditions in favor of a FLOSS model.

2.1 The FLOSS Competitive Advantage.

In a research conducted with Jean-Benoît Zimmermann and published in the European Management review (Jullien and Zimmermann, 2009), I proposed a model of duopoly competition between two for-profit firms, the first basing its offer on a proprietary software and the second on an FLOSS product. The ambition was to build a unified framework that could embrace both market and user feedback aspects for understanding why firms present such a wide diversity in their degree of involvement in FLOSS dynamics.

Our assumption was that, besides more usual arguments like the intensity of the competition and the extent to which software is a core competence of the firm, users' skills represent a major driver of this diversity. Users play a double role, deriving from both their economic and technical standing. Depending on the market, and especially their bargaining power in it, the users are more or less able to select the (technical) offers. At one extreme, users and contracts in the global service/architects market are related to large structures, with substantial buying capacities and generally endowed with significant technical skills. So they are likely to influence economic and technical choices. We distinguished three main types of users according to their relation to the product and the technology (Zimmermann, 1995a; Kogut and Metiu, 2001; von Hippel, 1988, 2002). The first is the category of "Naive customers or users" (that we denote N) who are not endowed with noticeable technical skills and do not individually weigh very much in economic terms. They are overall sensitive to prices and even if they may react to new characteristics of products or to branding issues, they are not capable to translate their needs and satisfaction into technical terms. The second is the category of "Kogut-Metiu Users" (KM)¹ who are not able to contribute to software development but can generate new features or innovations by revealing their own needs. KM users are sensitive to price and quality arguments. The third category is that of the "Von Hippel Users" (VH) who may act as "sources of innovation" (von Hippel, 1988, 1986) able to contribute to software development by proposing improvements or modifications, developing it by themselves or at least able to design the technical specifications.

In our model, strategic decisions are taken simultaneously, each agent formulating rational expectations about the behaviour of his competitor. We consider that the level of investment of each firm is decided while taking into account the expected skill level of the indifferent agent, thus of the respective market shares and the expected feedback effect for the FLOSS firm, as the more people are adopting the FLOSS solution, the stronger the community is the higher the quality of the FLOSS product. However, we modeled the strategic interaction of the two firms as a two-stage game, since there are two strategic variables for each company, the investment and the price. In the first stage, each firm chooses a level of investment that enables it to get an intrinsic quality level. The second stage was a price competition in which each firm chooses a price. This price had a recursive effect on the level of the indifferent agent, thus on the respective market shares and the level of the user feedback for the FLOSS firm, and thus on the level of quality it is actually reaching.

¹in reference to the notion of "frontier-users" put forward by Kogut and Metiu (2001).

Our results are the following: when the dominant user's skill is low enough, the FLOSS firm invests little, and stays at a lower level of quality than the proprietary one, targeting a relatively price-sensitive market. If the feedback effect is efficient enough, when the dominant user's skill is high enough, the FLOSS firm may invest a lot, obtain a higher level of quality than the proprietary firm, targeting a relatively quality sensitive market. The proprietary firm invests less, aims at a lower level of quality and targets a relatively less skilled and more price sensitive market segment.

Actually, the second result is in accordance with the work of Henkel (2006a), who showed in a duopoly model that "a regime with compulsory revealing can lead not only to higher profits, but also to higher product qualities than a proprietary regime".

Our model bridges those two extreme situations where the strategy of the FLOSS firms is respectively targeting either quality or price sensitive markets. It also echoes an old managerial choice that firms face when entering the computer market Cusumano (2004): either being a service company or a product company. Observable FLOSS strategies and our model both confirm what Cusumano proposed: 1) if there is a market for service, i.e. if customers are ready to pay for it (because they understand the added-value of this service), it is worth doing it. If not, firms must concentrate on selling quite standard products and rely on economies of scale. 2) There is a point where it is better to stop investing in services, because the costs are growing quicker than the market share or the willingness to pay. In between the two extreme cases, the opportunity for an FLOSS strategy does not vanish, but it is highly dependent on the technical nature of the product and the production process, and more particularly the possibility of exploiting a medium-skilled user base. One crucial aspect in this concern will be the importance of test tasks facing the development of the product. Kogut and Metiu (2001) show that these "frontier-users" who are not skilled enough to contribute by writing code are nevertheless capable of making a decisive contribution to software improvement, by constituting a huge test base and debugging base in a field where maintenance costs can reach 50 to 80% of the software budget. An interesting illustration is given by automated crash description programs included in most of the popular desktop FLOSS products (Firefox, Open Office).

This results conducted us to try to systematize the analysis of FLOSS firms' business model and the link between this model and their investment in the communities.

2.2 Firms' Strategies Regarding FLOSS.

The results we found via a mathematical model are very explainable. In a market where the dominant user is naive, firms can only differentiate on prices, as it is the only signal understood. The more competent users are, the more firms can differentiate vertically on quality or horizontally via the creation of niche products for specific needs and users. After explaining why FLOSS movement help solving certain recurring problems of the computer industry, and among them the respect of the norms, still with Jean-Benoît Zimmermann, in an article published in the *Revue d'Économie industrielle* (Jullien and Zimmermann, 2011b), I showed how the competences of the user may explain the industrial appropriation of FLOSS, and its variance from one market to another and the competitive advantage FLOSS gives. Here are the principal results.

2.2.1 Hardware.

This may be the sector where the link between structure of the market and FLOSS business practices is the clearest.

In the server market, producers have habitually provided proprietary solutions with proprietary Unix². The rise of PC servers has permitted some users to avoid such a bundling problem; moreover, using Linux or another free Unix means a cheaper offer (vertical advantage) reusing Unix programs (content) portfolio. Here suppliers are dealing with highly-skilled VH clients that have forced them to adopt FLOSS. The competitive advantage of FLOSS is its openness, allowing it to be tuned to specific needs, as proved by Google, which runs more than 40,000 servers with “customized versions of Linux”³.

On the segment of notebooks, where users are mainly naive, competition is overall based on prices. When Asus entered the market with its eee-PC, it used Linux for price reasons, because Microsoft Windows Vista was too costly in terms of resources needed and price to be competitive. Since then, considering the success of this market, Microsoft has designed a specific, downgraded version of Windows XP for these computers, and today, Linux market share has dropped to around 10%⁴. In that case free software is considered as a freeware and does not seem to provide any competitive advantage. It has only been used to force the monopoly to drop its prices.

²See West (2003) for a full discussion of FLOSS strategies in that sector .

³See <http://en.wikipedia.org/wiki/Googleplatform> for a summary of the Google platform and useful links on that topic and Stross (2008) for a whole presentation of Google.

⁴<http://www.i4u.com/article23707.html>

2.2.2 Software and service.

2.2.2.1 Software platforms.

We already showed that Linux had succeeded where users were VH. Of course for price reasons, but also for quality reasons: people could choose a system better adapted to their needs as it is more adaptable, as did Google. For a platform editor, the attractiveness of its platform is a growing function of the application available on it. The asset of a FLOSS platform is to allow the creation of a community of application providers, which will adapt their product to their platform. And, as these applications are not at the same level of maturity, since they may be incompatible, there is a need for an actor granting this compatibility, helping firms to select the applications they need. In a word, the “3A services” (assistance, assurance and adaptation to the use).

RedHat is the best example of such a strategy, with on one hand, the Fedora community to foster “innovation”⁵, the availability of applications on a Linux platform, and on the other hand, RedHat enterprise, when you want to buy assurance and assistance⁶.

2.2.2.2 Packages.

We said that a growing source of revenue comes from, again, “3A services”. Currently, the main evolution for those firms is to switch from a demand pull strategy (functionalities are developed to stimulate/create the demand) to an ‘on-demand’ development (development when required and paid for or carried out by the users). Therefore, FLOSS is used to increase the business feedback from users and by considering openness as a way to reduce transaction costs and as a signal of quality. This explains why open source business products are developed mainly in “business” software (ERP, computer infrastructure software like compilers), where users ready to pay for configuration, maintenance or assistance services are numerous. We distinguished two kind of structures, again regarding the computer skills of the users:

1. When users are VH software professionals, we find one firm organizing cooperation around its brand named product. The producer approves the contributions, ensures stability of the tool and helps developers to use it. If some individual contributor becomes important (in terms of contribution volume/quality/innovative aspect), s/he may be hired by a producer, with reduced recruitment

⁵See the presentation of Fedora at: <http://fedoraproject.org/wiki/Overview>

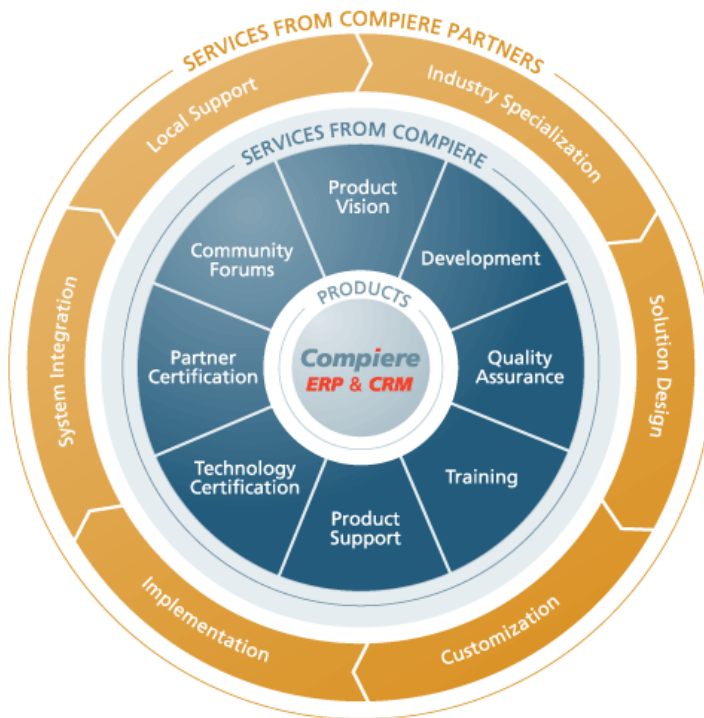
⁶Also very well explained on RedHat’s Web: <http://www.redhat.com/rhel/>

costs and risks (ACT or MySQL but also some small services companies are using this method). By contributing to innovation, the developers (and possibly companies using the tool), are therefore guaranteed that their needs will be taken into account more quickly and integrated into the product (which is a fundamental factor in reducing costs, according to von Hippel (1988)).

2. When users are more KM, firms are more service oriented. The open source asset is more in being a flexible, adaptable input, developed by a consortium of information system consultants sharing and co-developing the tools they are basing their business on, with sometimes a two-level organization, where a software producer and editor of a tool deals with information system resellers. The text-book examples are Compiere ERP and CRM, or Zope Content Management System (CMS), where there is an editing firm, which sells its services and products as in case one, but mainly to “partners”, service companies, as shown in figure 2.1 on the following page. In the second case, resellers act as service, architect companies we will discuss in the following paragraph. In both cases, the interest of FLOSS is its flexibility which is used to adapt itself to the client’s needs and the license, which guarantees that the product will always be available, as the modifications performed by other firms. the consortium organization, decreases the cost of development for each member (as it is shared) and facilitates and accelerates the achievement of a global presence for the product.

When naive users are dominant, it does not seem that firms manage to do a direct business based on FLOSS products. Even if some FLOSS offers exist for that segment, such as Open Office or Firefox, their market share remains small⁷. We do not know firms dedicated to them, and, according to us, firms’ support is more a consequence of platform or hardware providers strategies: these actors need these commodities for their platform to be adopted by VH or KM users, and FLOSS is a means to create a consortium to develop it (as SUN does by supporting Open Office development).

⁷Some sites estimate Firefox market share around 20% (see <http://marketshare.hitslink.com/browser-market-share.aspx?qprid=0>) but they probably over estimate it, as they look at the browser used to visit site, which thus favor the browsers used by Web intensive users, which are more skilled than the mean user, and may use more FLOSS.

Figure 2.1: The Compiere “ecosystem” (taken from: <http://www.compiere.com/partners/>)

2.2.2.3 Architects.

As pointed out, assembling components requires access to the source codes (problem of compatibility), and their adaptation to different needs (of users and other components) Horn. They must be available in the form of FLOSS software (therefore legally modifiable). The competitive advantage in using FLOSS, in addition to price, is therefore the ability to offer an assembled set of components with greater interoperability, which should increase the quality of the final product, on a market where the quality of services is one of the recurrent problems (De Bandt, 1995). Revenues are generated by assembling and adaptation services, as is the case for any traditional service company. The only uncertainty about the model concerns the availability of the components: who will develop them and who will maintain them? Moreover, the customers of these companies may already have (proprietary) programs installed that need to be taken into account. In the end, an open source strategy could even be a guarantee of means (maximum use of free software), but not a guarantee of the results (use of only free software), unless the customer requests this, since in this situation, s/he has the last word.

Two kinds of firms use FLOSS today: newcomers who specialize in FLOSS architecture, using FLOSS as a vertical (price) and horizontal differentiation asset⁸, and incumbents, such as IBM for its service activities. Traditional service firms like Cap Gemini are more agnostic with regard to the technologies used and the intellectual property regime involved. They will generally follow the customers' demand which depends on their ability to keep up with the development of the project. These customers are most often large organizations, skilled computer users that are receptive to the opportunity to integrate the most advanced software components, developed under FLOSS licenses. So they are becoming increasingly involved in FLOSS as the market grows and matures⁹. Global service firms' Web site are quite explicit on this strategy¹⁰.

2.2.3 The dependence on users' skill.

Table 2.1 summarizes our findings.

What seems clear after this rather qualitative analysis is that the skill of the users matters for understanding the level of implication of firms in FLOSS. When users are naive, firms may use FLOSS, but only for price reasons, in the same way as they could use freeware. The more VH users are, the more complex strategies regarding FLOSS are, and the greater firms' involvement and participation. In some cases, when users are VH, firms may even produce FLOSS and animate the community like Ada Core Technology for Ada 2005 and MySQL AB for MySQL data bases do. But in any case, FLOSS is regarded as open source software. This means that in that case, firms use FLOSS for technical reasons (sustainability, flexibility) and for innovative reasons (increasing the speed and quality of feedback).

2.3 The Role of Firms in FLOSS Communities.

Is the consequence of firms market position that they will behave differently according to the significance of FLOSS software or FLOSS community in the specificity of its offer: a complementary or a core asset for this offer? According to the theory (Teece, 1986), if yes, firm should invest a lot to manage this asset, if

⁸As explained by Slatter (1992), one of the main strategy for newcomers in technological market is technological differentiation. Basing its offer on new FLOSS products can be seen as a way for new service companies to differentiate.

⁹In 2005, Gartner forecasted that "[in] 2008, 95 percent of Global 2000 organizations will have formal FLOSS acquisition and management strategies" (<http://www.gartner.com/it/page.jsp?id=492152>). In their 2008 study, they said that "Adoption of open-source software (OSS) is becoming pervasive, with 85 percent of companies surveyed currently using OSS in their enterprises and the remaining 15 percent expecting to in the next 12 month" (<http://www.gartner.com/it/page.jsp?id=801412>).

¹⁰See Capgemini's for instance: <http://www.us.capgemini.com/services/servoverview.asp?ServID=30>

Table 2.1: Firms' FLOSS strategy according to the skill of the users.

	Naive	Kogut-Metiu	von Hippel
Hardware	FLOSS as a Freeware (Linux, with Open Office and Firefox). Price differentiation. Ex.: Asus' EeePC.	FLOSS to create a consortium on commodities development. No business. Ex.: Open Office and Firefox are installed in computer to increase the applications available for the same price.	FLOSS as technical tool increasing the quality over price ratio via horizontal (adaptation to the needs) and vertical (cheaper use) differentiation. Ex.: PC servers running Linux or *BSD.
Platforms	No FLOSS strategy	FLOSS to create a consortium on commodities development. No business. Ex.: Open Office and Firefox.	FLOSS to foster applications availability. Firms sell selection, compatibility setting and stability over time. Ex.: RedHat
Business software	No FLOSS strategy	FLOSS to create a consortium of consulting firms selling adaptation of the product.	FLOSS to create a closer relationship with users (quicker and better feedbacks). Firms sell 3A services. Ex.: MySQL or Ada Core Technology.
Service	FLOSS to propose cheaper services? No example known.	FLOSS as a flexible commodity to propose/sell cheaper service solutions. Ex.: Compiere ecosystem.	FLOSS as a flexible commodity to meet users' demand and to propose more flexible solutions. Ex.: Cap Gemini, IBM global Service

not, it should not invest at all, buying or using it as it is (as a component “of the shelves”). In an article published in the Journal of Innovation Economics, Jean-Benoît Zimmermann and I showed that firms seem to behave as so (Jullien and Zimmermann, 2011a).

2.3.1 Naïve users, the freeware strategy. The community as a commodity.

As, in that case, the aim of using FLOSS is to propose the lowest price possible, firms will not invest in FLOSS development more than the effort needed to adapt the software to their product(s): this would increase their cost. So FLOSS may be seen as a free commodity, a freeware.

Of course, one might put forward that such strategies have a flavor of free-riding and the risk is to demotivate the most committed people in the communities¹¹. However, by adopting FLOSS products, those firms participate in the expansion of the FLOSS users' network. In a competition regime in which the battle for network externalities and standards play a crucial role, this may be considered as a strong boost to FLOSS, that can at least win the community's neutrality if not approval.

¹¹See Foray et al. (2007) for a discussion of that particular point.

2.3.2 Business packages for VH users. The community as a core asset.

As far as business packages are concerned, the specific asset of the producer lies in its package knowledge and in its capacity to manage the dynamics of evolution. This makes the open sourcing of a software the specific asset of the firm which owns it: on the technology markets where the customers are computing developers, revealing the code facilitates cooperation. The producer organizes the collaboration in a “symbiotic” relationship (using the terms of Dahlander and Magnusson (2005)). Developers (possibly companies using the tool), by providing their own innovations, are thereby assured that their needs will be taken into account more rapidly and integrated into the product, a crucial point to reduce their costs (von Hippel, 1988); from the producer’s point of view, this decrease the R&D cost as the users provide him/her with new feature requirements and, more original, implementation; on the other hand, only the one who integrates contributions is capable of verifying and of guaranteeing their correct functioning and to help clients to use it. So, a FLOSS based package model means that the firms which publish the software remain heavily involved in its development in order to control it. As their core competence lies on the management of the software edited, the companies should only invest in the software they edit, and the involvement of salaried developers in other projects should not be encouraged.

2.3.3 Services to VH and KM users. The community as a complementary asset.

In between these two extreme cases of involvement into communities are service firms. The only uncertainty in their FLOSS business model arises from the availability and the quality of FLOSS components: who develop(s) them, who maintain(s) them? So they need to evaluate these components and to monitor their evolution¹². This need for evaluation and control increasingly depends on the importance of the component for their business and that of their clients. We may even formulate the hypothesis that the more skilled the users are, the more the firm must master the technology, because of the growing level of complexity of the feedbacks and demand.

And to be able to integrate knowledge and innovation from the open-source communities, open-source firms have to develop internally efficient capabilities of absorption, an essential condition to capitalize and

¹²Considering the evaluation part, it is worth noting that the main service companies in France have published tools to evaluate FLOSS, and they use this as a commercial argument (we would say a signal) to their clients on their capacity of evaluating these products. See, for Atos, <http://www.uk.atosorigin.com/en-uk/services/solutions/systemsintegration/technologiesexpertise/opensource/default.htm>, and for Cap Gemini, <http://www.capgemini.com/services/technology-services/open-source/solutions/>.

internalize the communities' contribution and the users' feedbacks to improve their own product quality. Dahlander and Magnusson (2008) working on the relations between firms and open-source communities show that those firms need "to develop sufficient absorptive capacity to benefit from external developments, not only to identify useful external knowledge, but also to assimilate and apply it". This is what has been called a "commensalistic approach" (Dahlander and Magnusson, 2005). This corresponds to the more general assertion from Cohen and Levinthal (1989, 1990) about the necessity for a firm to make internal efforts of R&D a prerequisite for the absorption of external technology.

We consider this reflects a change in the technologies used, thus of the complementary assets these firms need to manage, not really in the core competences. Traditional architect firms are not involve in FLOSS development, as they do not use these technologies. But they may have other processes for monitoring the evolution of the complementary asset, the technologies they use. They may participate in editors' training sessions, or conclude "global alliance" with their key partners, as Cap Gemini does¹³.

2.3.4 Conclusion.

These results are consistent with Teece's theory (Teece, 1986; Teece et al., 1997), if considering the FLOSS community and the evolution of a technology as the asset (see table 2.2). If firms see it as a source of technology they do not invest. They do invest a lot if it is the core of their business to guarantee a efficient integration of the innovations coming from the community to this community. In between these two extreme cases, the community is seen as a complementary asset, as firms try to participate to follow and control the innovations coming from the users.

Thus firms may undertake the different "roles" of the FLOSS organization onion model¹⁴ (figure 2.2 on the next page): core developer (like MySQL), developer (IBM with Apache), bugs fixers or reporter (Compiere "partner" companies) or user (Asus), and this mainly depends on their users.

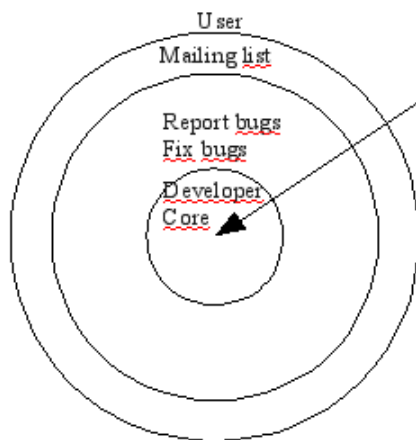
Quite logically, if the involvement of the market into the production of open, online communities is that dependent of VH users, something which seems to be validated by the models of competition (Darmon and Torre, 2009; Jullien and Zimmermann, 2009; Darmon et al., 2011), the stability of such

¹³<http://www.capgemini.com/collaboration/alliancepartners/> for classical alliance, and <http://searchsystemschannel.techtarget.com/news/article/0,289142,sid99gci1261207,00.html> with those done with open source world.

¹⁴This model is described by Herraiz et al. (2006) and has been proposed by Crowston and Howison (2005). In a career in a community, the successive tasks a developer may do are to use the program, to use the mailing list(s), to report bugs and to fix them, to be a core developer. See section 3.2, page 41.

Figure 2.2: Onion model based level of involvement in FLOSS communities targeted by firms regarding their business and the skill of their users.

(a) Onion model of the level of involvement in FLOSS communities (from Herraiz et al., 2006)



(b) Level of users' skill and level of involvement of firms in FLOSS communities.

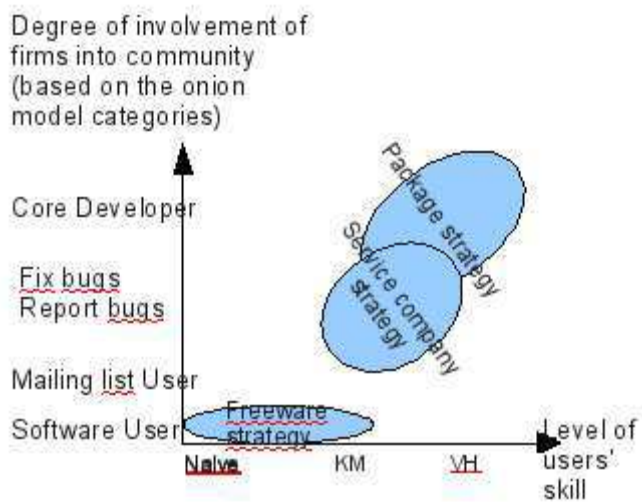


Table 2.2: Firms' investment in the asset "community".

	single firm community	consortium	open community
Package firm	one dominant	contribute marginally if needed (to make its package work correctly)	contribute marginally if needed (to make its package work correctly)
Service firm	from technology taker to marginal contribution in function of the importance of the technology	from technology taker to marginal contribution in function of the importance of the technology	from technology taker to marginal contribution in function of the importance of the technology
Freeware strategy	technology taker	technology taker	technology taker

communities depends on their capacity in attracting such users, and in making them participate. That is the reason why, in a second part of my research, I tried to better understand the motivations and the behavior of the community of creation participants.

Chapter 3

Individual Involvement in Communities of Creation.

This chapter deals with the question of involvement and recruitment of people in communities of creation: what are the basic motivation for participating? are the future big participants detectable soon enough to be welcomed and nested to improve the recruitment efficiency? are some communities more efficient than others to do so, and why?

Regarding the reason why people get involved in communities of creation, and beyond the user-as-innovator incentive, theoretical analyses of incentives, in software projects (Foray and Zimmermann, 2001; Lerner and Tirole, 2002) or in wikis (Forte and Bruckman (2005), using Latour and Woolgar's (1979) analysis of science "cycles of credit"), estimate that the other main vector for participation is the quest for reputation. Applied works on Wikipedia (Nov, 2007; Yang and Lai, 2010; Zhang and Zhu, 2011), electronic networks of professionals (Wasko and Faraj, 2005; Jullien et al., 2011), and open source software (Shah, 2006; Scacchi, 2007), confirm that peer recognition, whether it be professional or community recognition, is a main motive for participation, in addition to intrinsic factors (personal enjoyment and satisfaction from helping by sharing their knowledge). This argument that on-line volunteer participation can be explained by the same incentives found in science has also been used in reverse. Schweik (2006) uses the same framework by Hess and Ostrom (2006b) to describe open source organization in order to extract its main mechanisms to construct an open science project. But can we imagine that these coming-after motivations, these motivations for the most involved, are sufficient to make people involved

in the production? The lack of research on the subject suggested to us a qualitative approach to develop initial, but detailed, exploratory results (Von Krogh et al., 2003). In a work done with Karine Roudaut and published in *Management International* Jullien and Roudaut (2012), I developed this approach via interviews with researchers in algorithm-based industries. Our study question, to quote Yin (2009), has been to understand if software producers may find interest in participating in open source projects when they are not users, and why.

Secondly, if the motivations change over time, this means that they are discovered by the participants and, from the viewpoint of the project, makes it difficult to know who is going to really involve himself in the project. This process of joining is constantly needed, first because there is a decrease in involvement with the number of years people are active in the community (Ortega and Izquierdo-Cortazar, 2009 for open online communities, Borzillo et al., 2011 for intra-organization ones) and secondly because the production of knowledge needs a constant renewal of ideas (Arora and Gambardella, 1994). Specific programs, targeting newcomers, have even been designed to mentor them (see Mateos-Garcia and Steinmueller, 2008 for open-source project Debian, and Musicant et al., 2011, for Wikipedia), and a lot of studies have been conducted to better understand the motivations for contributing (von Krogh et al., 2012 proposed a review of these motivations and of the literature studying them), as the different steps to join the core (Von Krogh et al., 2003; Hertel et al., 2003, and more recently Crowston, 2011). If these studies are interesting to understand why the contributors involve themselves, or stay, they are of little help to the core members, the mentors of such communities in identifying the promising participants, and in guiding those beginners toward sustainable participation, without putting too much pressure on the less invested (or less investing) people. The division of labor and the organization of this division in such open communities of creation has been discussed by Rullani and Haefliger (2013), opening a discussion on how to manage those different contributors. They stress the importance of “caring about the periphery” and “not trying to drag its members to the core”, as not all the members of this periphery are destined to become core members, and as putting too much pressure on them may even prevent those people to keep on contributing (Halfaker et al., 2011). Behind this is the argument that participation is to be learned via progressive involvement and increasingly complex contribution and interaction with the members of the community, and refers to the progressive involvement in communities of practices theorized by as “legitimate peripheral participation” (Lave and Wenger, 1991). In an article written with Sylvain Dejean (Dejean and Jullien, 2012), I wonder if it is possible to identify the different contributors early in

the process of joining, early enough to adapt the caring to the producer's profile and thus decreasing the discouragement rate of both newcomers and sponsors. In other words, I was looking for what people need (in terms of competences and help) and what they do (characteristics of the contribution) to cross the invisible frontier between the users and the providers, to paraphrase Ostrom, and if the circumstance of this crossing informs about the level of contribution those participants will reach. We will focus more on the core participants, as, if both core and peripheral ones are needed (O'Mahony and Bechky, 2008; Hemetsberger and Reinhardt, 2009), as in most collective actions and projects (Ostrom, 1990), the first is much rarer than the second, in addition to being more productive (Lakhani and von Hippel, 2003; Mailart et al., 2008; Voss, 2005; Kittur et al., 2007a). This estimation is done via econometrics and Probit models (with Heckman correction to account for potential selection bias in our sample, Heckman, 1979) on a sample of more than 13,000 Wikipedia users, and sometime contributors (we explain below why we choose Wikipedia).

Third comes the question of the efficiency of the community in providing enough incentive for the persons to get involved in. There is a growing concern about the difficulty of recruiting and retaining new editors (Halfaker et al., 2011), leading to the development of specific programs (Musicant et al., 2011). This seems to be a recurrent problem in open communities, already stressed and studied by Von Krogh et al. (2003) in the case of open source software community joining. There are many possible explanations for this situation, but the literature on open source software projects (Koch, 2008a), and on collective action more generally (Marwell and Oliver, 1993), suggests that such a slow down may simply be the result of the project entering a mature phase in which it needs fewer additions and thus fewer contributors. However, a more troubling possibility is that the evolution of Wikipedia has led to the development of processes that make contributing to these projects more difficult, making the work less rewarding (Ransbotham and Kane, 2011), raising invisible barriers to participation for outsiders (to take Ostrom's perspective). With Kevin Crowston and Felipe Ortega, I tried to assess this efficiency in the case of Wikipedia. We chose Wikipedia because, for each language, there is a separate version of the encyclopedia, a different project, with its own editor community and collection of articles. Importantly, the projects are at different levels of maturity, some quite mature, others still getting started and others somewhere in between. However, they all share the same tool for collaborative edition (MediaWiki) and the same basic rules for collaboration, the "five pillars" of Wikipedia¹. As well, the global structure of the

¹<http://en.wikipedia.org/wiki/Wikipedia:Fivepillars>.

projects, measured as a network, the nodes being the articles and the links the links between the articles, seems to be about the same, at least for the main projects (Zlatić and Stefanić, 2011). In contrast to studies on open source software (see for instance Crowston et al., 2006; Koch, 2009) that compare project that use various technologies, programming languages and collaborative tools, this uniformity may help us to better understand, in their difference, what differences are due to process evolution.

3.1 Beyond the Private-Collective Innovation Model?

Discussing von Hippel and von Krogh's argument that these online open communities are based on the concept of user-as-innovator requires to first analyze the problems raised by this division, before identifying an industry where these problems may be overcome, where the conditions are the most favorable to finding successful projects. Still considering Hess and Ostrom's framework (2006a), splitting the user-as-innovator means to restudy the action arena, or how people interact to define and produce what is needed (what we call the "FLOSS factory"), and the cost-benefit for the providers to participate in this project (the "FLOSS incentive regime"), to go to the action arena.

Considering the fact that the most cited and found motivation to participate in such a kind of project is the reputation, the chance to find a successful FLOSS production project where providers are not users should be enhanced if the producers belong to the scientific community, where reputation is one of the main rewards (Foray and Cassier, 2001b,a). This should be the case, even if the sociology of sciences (Lamy and Shinn, 2006a) shows that the commitment to non-academic projects (in that case entrepreneurship) may be weighted by potential concerns about the "scientific value" of such non-traditional academic production, the "attachment" to the cycles of academic valuation.

As I said, the lack of research on the subject suggested us a qualitative approach to develop initial, but detailed, exploratory results (Von Krogh et al., 2003), and we developed this approach via interviews with researchers in algorithm-based industries. Our study question, to quote Yin (2009), has been to understand if software producers may find interest in participating in open source projects when they are not users, and why. What is common to these fields is that designing a "better" algorithm requires competencies in statistics and classification, a form of applied mathematics. These are not integrated into biology, earth observation or communication, but rather are autonomous disciplines referenced by IEEE. This separation was the insight which made us think about a distinction between algorithm producers and

users, and was confirmed by the study. The second reason for studying data processing for knowledge-based sciences was that there were FLOSS initiatives, more or less visible at an institutional level. Factors influencing participation.

3.1.1 The construction of the absorptive capacities.

As explained by OECD (1996), the way and conditions in which an algorithm can be used are complex and require specific skills (“know-how”). In other words, even if the conditions of use of the algorithm are codified (described) and a program exists for making it run on a computer, this is not always enough for the end user to be able to use it correctly. On the other hand, users express needs that providers have to understand, and translate into software specifications. This is what Zimmermann (1995a) called the “technologies of use” (our translation from French). In other words, both knowledge users and providers have to develop their “absorptive capacities”, or their capacity for understanding what others produce, in order to decrease the “tacitness” of the knowledge and to pass it across borders. In traditional “user-as-innovator” FLOSS organization, and, as already mentioned, these absorptive capacities are less an issue because, per definition, the users directly translate and produce, even if, as Jullien and Zimmermann (2006c) have shown, this is not completely solved for those users who are not providers.

Social psychology and computer use studies have brought some answers on how users and producers construct the absorptive capacities²: bridges have to be established by the design of specific “boundary objects” (Star and Griesemer, 1989). The goal is to implement the knowledge in an “appropriable” way for the user, or/and by specific “frontier persons” coming from one side of the boundary but having developed competencies on the other side. These boundary spanners or “intermediaries” have been described by Actor Network Theory (ANT), or the Sociology of Translation (Callon et al., 1986; Akrich et al., 2006), as “mediators”, a role which involves a “translation”, here between two distinct “communities” of research, involved in a collective action. From the organizational point of view, because of his position, this actor is close to figure of the “marginal sécant” (influential outsider), as described by Crozier and Friedberg (1997). Barcellini et al. (2008), for the open source project Python, and Laniado and Tasso (2011), for Wikipedia, shown the importance of these actors for the functioning of open knowledge communities.

²For a review of the literature and a presentation of the case of bio-computing, see O’Day et al. (2001).

In addition, an algorithm (as defining and defined as a “piece of knowledge”) does not work alone, but has to interact with others. This is the “modularity” issue, i.e. “how parts are grouped together and about how groups of parts interact and communicate with one another” (Koppl and Langlois, 2001, p. 18). This articulation of basic pieces of knowledge to create more complex knowledge exists in software (Langlois and Robertson, 1992), molecules (Bureth et al., 2007) and chains of treatment (coding-decoding of a voice signal for telecommunications, gene sequencing and analysis in biology, medical image processing, etc.) One has to make sure there is no incompatibility between the different elements of the chain, that what the “out” data produces is what is expected as “in” by the following element in the process.

So to define the investment producers have to make, we had to understand how the interactions are constructed. This will be our first preliminary study question. We formulate the hypothesis that a FLOSS factory may help for the construction of such interactions because, as Bessen (2005) explained, it clarifies the interfaces between the different components (thus decreasing the part of tacit knowledge), facilitating producers’ work, as they then only have to focus on the production of a single piece of knowledge) as well as that of intermediaries (i.e., the creation of chains of treatment is facilitated).

3.1.2 The projects’ environment.

Finally, looking at FLOSS producers participation to projects may disguise the importance of the institutional approach in these projects. Producing open source algorithms may be, for the sponsors of the platform (institutions of services within an institution dedicated to computer support for researchers), a way to facilitate the coordination and the construction of standards (Simcoe, 2006), because they depend on these standards. They use them for their own production, either as user or as producer of technology using software bundled with technological outputs (such as data acquisition and processing offers like remote sensing tools). Olson (1965) argued that there may be transfers to ensure participation in collective good, and that institutions are needed to organize these transfers. But just because there is interest in FLOSS projects at the institutional level does not mean that, at the individual level, the transfers or institutional support are important enough to support producers’ participation. So, in our case, the structure supporting the platform project (service, contributors’ reward) may appear as a factor explaining (non-)participation, as already noted when we referred to Lamy and Shinn (2006a), but it can not be our point

of entry in this study.

3.1.3 Data and method.

Data collection was performed between the end of 2008 and the end of 2009 and consisted of interviews (21 semi-guided interviews of more than 90 minutes each). All the interviewees were scientific professionals (researchers and research engineers), 18 belonging to public institutions (institutes of technology, CNRS, Institut Pasteur, CNES) and three to firms or private institutions. As the FLOSS movement has its roots “in the university and research environment” (Bonaccorsi and Rossi, 2003, p. 6), one can argue that these producers should be more open to such arrangements³. We had to collect two types of points of view regarding the production of algorithms and their diffusion in the chosen disciplines: the algorithm producers and the producers of the platforms.

What we wanted to empirically appreciate was the participation of the researchers in FLOSS production and how this production is organized in a FLOSS factory. For this purpose, we collected information on the activities and production of the researchers and research engineers in the selected disciplines (bio-computing, remote sensing and digital transmission). We interviewed people about their representations of their production in their scientific environment and about the definition of what a “good algorithm” is. We also interviewed them on the existence of joint work, namely collaboration practices between the different actors contributing to and using the chains of treatment. Amongst the interviewees, some participate in the cooperative development of platforms of knowledge production (some open source, others not), while others do not. This allowed us to identify why they do or do not participate. We stopped the collection of new interviews when the exploration of the content of each new interview did not bring additional significant meaning (a summary of the methodology is available in table 3.1, page 39).

While the total number of interviews may appear low, their length allowed us to collect a fair amount of rich material (more detail and variance), and to identify some “coherence in attitude and social behavior” embedded in “a historical path, both personal and collective”, to quote and translate Beaud (1996), the result expected from this kind of qualitative analysis. “This path is personal, as each interview describes the trajectory of a scientific actor, but also collective because it describes the specific scientific field it is embedded in” (ibid). According to Flyvbjerg (2011, pp. 301-316), referring to the definition

³See, for instance, the open source science initiative, <http://www.opensourcescience.net/>, or the online initiative to cure tropical diseases, in biotech, <http://pubs.acs.org/cen/science/84/8430sci1.html>.

Table 3.1: Research methodology and interview guide.

Choice of the methodology	Qualitative research approach by means of semi-structured interviews with individual actors participating (or not) in collaborative platform projects (FLOSS factories). Methodology based on Beaud and Weber (2003); Blanchet and Gotman (2010); Bertaux (2010). We identified the players interviewed given a) their visibility in the selected collaborative platform projects, or b) their visibility in a disciplinary field in which the selected projects are developed. In our investigation, we have always given priority to the individual (personal and professional) dimension of the contribution to these projects against the institutional dimension.
Data collection	Development of an interview guide based on the one used in a research project on on-line communities of practice (ANR CCCP-Prosodie). Semi-structured interviews (21) with researchers and research engineers in the disciplines selected, contributing (or not) to these platform projects. Recording and full transcript of the interviews.
Interview guide. Adapted from http://www.cccp-prosodie.org/spip.php?article40	From interviews, we wanted to include the participation, or lack of participation, in this type of project, the terms of participation (eg. if it was part time or not), and a commitment (or not) to the the idea defended by these platforms in the scientific production. The main themes developed in the guide were: <ul style="list-style-type: none"> • Description of who the actor is, his job, his activities. • Description of what he does in the selected project, the way he contributes to the project (production, how he works and produces knowledge, role, relation to the production, relation to the project). • Description of the project's activities (relations, actors, activities). • Commitment to the project: description of the process steps of the actor's commitment to the project, from the first contribution to greater involvement. The goal is to identify the various registers of the commitment. Evolution of the actor in the career of the project. • Evaluation by the interviewee of his project participation (relative to the project itself, the job, and the scientific field). • Professional career.
Data analysis	Classical sociological manual procedure to analyze thematic content of the interviews: a horizontal analysis, i. e., for each interview the characterization of the main themes emerging from the discourses, then a transverse or vertical analysis for all interviews, i. e. the identification of the common themes amongst the interviews, and the confrontation of the actors' positions on each theme. Content analysis on each selected disciplines, then on all the software platform projects.

of Merriam-Webster's Online Dictionary (Merriam-Webster, 2009), "case study focus on an "individual unit", what Stake (2008) calls a "functioning specific" or "bounded system".[...] Finally, case studies focus on "relation to environment"[...]"'. In our case, the emphasis is on analyzing the actors' relationships with their scientific knowledge production, and their environment.

3.1.4 Results.

All the actors we interviewed agreed on the aim of the collaboration: proposing the best methods to extract “pertinent” information from physical data, using algorithms. However, this “best” does not mean the same thing to everyone: the algorithm producer would look at mathematics lock-ins to be solved, while the user would look at the quickest way to solve a problem (not always the most efficient), either looking at an already implemented algorithm to do the job or, when this is not possible, taking it to an algorithm producer who can understand them. This means that they have to invest time and money to understand each other, to develop “patterns of interaction”. Actually this investment seems to be made more by specific actors (the “boundary spanners”) than by the use of tools such as the open source platforms developed, in the project we studied. Non-use is not due to an a priori discrimination against FLOSS, but rather because these projects seem to provide less help to boundary spanners and few incentives to algorithm producers (in terms of reputation or of institutional incentive), whereas they have important extra investments to make the programs they developed open source.

All the researchers we met are open-minded regarding FLOSS. They may use FLOSS-based computers (with GNU/Linux operating system, or/and Firefox browser) and some disciplinary FLOSS programs, which can “facilitate their work” because they do not have to “redevelop standard applications”. Most of them (both algorithmists and end-users) think that, at a global level, having access to FLOSS programs implementing the algorithm would facilitate access to “knowledge” (of the existence of a new algorithm, of its performance, of how to use it, etc.)

But, as far as algorithm producers are concerned, even if they have a positive opinion of open source, open-sourcing the program they have developed when conceiving a new algorithm (to test it) is not obvious. This weak appetite for publishing FLOSS programs should also be partially explained by the fact that the gains expected do not cover this extra investment (it is not just a question of publishing the program they have “cobbled together” for their own need, under a free, open source license. On the other hand, access to new algorithms is facilitated by their open-sourcing, as it accelerates the appropriation of new algorithms. It is not a necessity as these users understand the methods published in scientific journals and are able to re-develop the implementation of the method, whether it is software (bio-computing, remote sensing) or hardware (digital communications), but it is a real added value. To sum up, we are facing a classic collective action paradox: even if agents agree on the fact that a FLOSS organization will improve

the efficiency of the domain, on both the algorithmist side and the end-user side, no one seems to be ready to invest the extra cost for producing the IP.

This study of “applied” science knowledge production has shown that today, FLOSS is not a solution on its own to knowledge diffusion, because there are not enough incentives for researchers to publish their software, and there is a need for extra investments to integrate the software produced into a chain of production. The design of standard platforms may help for that, but for the time being, knowledge producers still don’t have incentives to contribute to these platforms. When users are not producers, in contrast to the traditional FLOSS incentive regime supporting a FLOSS factory, the other two traditional regimes do not seem able to take over in the long run. This situation may change with the evolution of the system of evaluation in science, which has been initiated in bio-computing.

What this work shows is that even if there is a common goal to focus algorithm production on more efficiency, productivity and re-usability, and a common agreement to say that FLOSS may be a tool to do so, the perceived rewards are too low for both users and producers for them to invest in the extra cost of involving themselves in the project.

We may not have looked at the right platforms, because they are still in their infancy. But we think this works in our favor: when the developers are not the users, the initial phase to build a platform, which is the harder phase in a collective action (Marwell and Oliver, 1993), is too difficult to overcome to create a sustainable project, because people do not anticipate the success of the platform and the additional reward it may provide. The financial incentive may be used to bootstrap the platform with content, but, even then, it did not appear sufficient to attain the diffusion phase. Before building a new “open source” garbage can, the sponsors or the managers of such projects have to create sustainable incentives, which may be connected with people’s work: here it means scientific reward or obligation (you have to open source the code used to be published, for example). In other fields, these rewards remain to be defined and may be field-specific.

3.2 Individual Profiles of Participation.

Taking the conclusion of the precedent section, i.e. that it seems that the producers have to be users, as a starting point, questions the way people get involved in those communities, and if this investment is made by a certain kind of user. This question meets a growing concern of community managers about

the difficulty of recruiting and retaining new editors, a problem already noted by researchers (Ortega, 2009; Halfaker et al., 2011), leading to the development of specific programs for Wikipedia, for instance (Musicant et al., 2011). This seems to be a recurrent problem in open communities, already stressed and studied by Von Krogh et al. (2003) in the case of open source software community joining. Wikipedia, as other online communities, is socio-technical project (Bryant et al., 2005; Benker and Nissenbaum, 2006), where the tools and the rules mediate and shape users' activity around open collaborative writing.

Open source communities, which have been the case studies of lots of articles regarding the involvement (Rullani and Haefliger, 2013, von Krogh et al. 2012) present a particularity regarding the discussion on the different steps between user and core developer: each project is a combination of a (or several) community of practice (how to use the software), relying mainly on forums, mailing lists, and an epistemic community in charge of producing the software, the knowledge (Cohendet et al., 2001). In communities of practice, situated learning and identity construction terms are formally linked to the transmission of the codes and of the existing set of knowledge from mentors to newcomers; in epistemic communities, the exchange and interaction is about the construction of (new) knowledge formation and people are evaluated on their capacity to produce this knowledge (Amin and Roberts, 2008). Regarding open online communities, the first arena is about how to install, use the software, how to cope with specific problems, etc. the second is about the number of contribution, the quality of the contribution, etc.

As already said, there are bridges between those two arenas (Rullani and Haefliger, 2013 for a discussion of these bridges, Barcellini et al., 2008 for a case study of who do the bridging in the case of the open source language Python), and there are some kinds of learning in the epistemic community (Bryant et al., 2005; Mateos-Garcia and Steinmueller, 2008): interacting with core developers on the production of knowledge, new contributors learn how the production is structured and what is needed and are identified by those central persons. In the interaction, the motivations (the reasons why people contribute) change too: Shah showed, in the case of open source, that long-term participants enjoyed programming and interacting with the rest of the community (i.e., labeled as "hobbyists"), whereas short-term participants were typically driven by an immediate need for software (i.e., use value); if, for the most involved in Wikipedia, the recognition from the peers ('credit') is an important motivation (Forte and Bruckman, 2005; Bryant et al., 2005), as is the sense of mission (Liang et al., 2008; Prasarnphanich and Wagner, 2009), for most of the (small) contributors, the will to fix mistake is the principal motivation, making these people not strongly committed to the project (Kamata et al., 2010). According to Shah (2006),

this echoes the more general sociological notion of “career” (Becker, 1960, 1963), which stresses that people’s motivations and actions are curved by the social interactions they meet in their practice.

But whether people become very involved because they are more sensitive to those global motivation, the system being a way to identify them, or whether people integrate those global motivations (with other global rewards) interacting, remain matter of debate and investigation (von Krogh et al., 2012). Idiosyncratic social skills such as conversational knowledge are probably an asset to get involved and recognized by the community as “the probability that students with a high level of conversational knowledge leave [the Google Summer of Code project] is 64.5% lower” (Schilling et al., 2012). Psychological characteristics, such as “agreeableness, openness, or conscientiousness”, seem also to matter (Amichai-Hamburger et al., 2008). But all these characteristics are very hard to identify by the active members of the project, who are central for the promotion of a contributor from peripheral to core. Fang and Neufeld (2009) did not find evidence of a link between the period of apprenticeship as a user and the fact of becoming a core contributor in open source project either, as some people directly enter the project as central developers. This may be explained by the fact that they are hired from another project, where they did this period, or by the fact that learning to contribute can mainly be learn offline reading the documents and/or the discussions, what Edwards (2001) called the “private sphere”. And, according to the projects⁴, and to the theory of epistemic virtual communities (Amin and Roberts, 2008), the recognition in epistemic communities comes with the participation in the production of pieces of knowledge (the provision of code for FLOSS, the work on the articles, or the “edits” for Wikipedia), not firstly because of social characteristics or participation in discussions.

This makes me argue that the production of new pieces of knowledge is the real beginning of a sustainable involvement in open online communities and is the entry point for the project (and its regular contributors) to inform on the capacity of newcomers to involve themselves in the ‘collective sphere’, still using Edward’s words (2001), to become active members of the project. We argue that the private investment needed to do this first contribution is so high that those whose cross this border will invest deeply in the community. Considering the discussion below, some social characteristics, like social connections (knowing participants in the community, being mentored) may decrease this cost and facilitate

⁴Regarding Wikipedia, when projects have rules for running for administrator, they are about knowing the rules, but also about the number of edits (more than 3,000 and more of one year of activity for the French Wikipedia, https://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Candidature_au_statut_d'administrateur). O’Mahony and Ferraro (2007, part II), on Open Source project, showed that “developers who were making greater technical contributions (in terms of impact but not effort) and who were more engaged in organization building were more likely to become members of the leadership team”. (p. 1096)

the involvement, some technical skills (programming) may do too, and will have to be controlled.

Regarding Wikipedia, Bryant et al. (2005) studied the trajectories of involvement of some very involved participants, and showed a transformation of the goals and the activities of these participants during this phase: they evolve from self-oriented contributions and justifications to community-oriented contributions and justifications. This suggests that the difference between the level of involvement can be explained by what happens during this period. In the same time, for these most involved Wikipedians, the level of participation seems strongly dependent on the first contributions to the project as a registered user (Panciera et al., 2009). But when people registered in Wikipedia (or in another community), they may have spent a lot of time participating to this community as user, or even contributing anonymously, and Panciera et al. (2009)'s study would only reveal what happens at the end of the apprenticeship phase, as "another convention that is understood by Wikipedians but not by novices is that anonymous contributions are inherently suspect, so new users are encouraged to register and get usernames and to always sign their contributions to discussions" (Bryant et al., 2005, p. 8). A long period between the discovery of Wikipedia as a reader and the first contribution would suggest the existence of an observation period, an apprenticeship process. Our results suggest that, as only 39% of the contributors were registered when they did their first contribution.

In a research conducted with Sylvain Dejean, we surveyed the users of the French Wikipedia project in 2011, and analyzed the more than 13,000 responses obtained via econometric models (Probit and Heckman models) to assess what the first contribution may teach about the future level of involvement in the community (Dejean and Jullien, 2012). More precisely, our goal was to estimate the probability of becoming a big or at least regular contributor in Wikipedia as regard to these first steps in the community. Many information are contained in the time, the form and motivations underlying first contribution. Studying if there exists different paths to integrate the project, or on the contrary if people engage in contributing activity since the beginning of their discovery of Wikipedia should provide with a better comprehension of the users' involvement paths and of the organization of the open collaborative communities. We present here our main result after a quick detour to look at the methodology used, as I think I gives some insights about how to survey an online community.

3.2.1 Surveying a community of creation. Data and method.

Online communities produce complete and available data on the contributions, but, unfortunately few information on the participants. Wikipedia is a good example of this fact, as information on contributors regarding their skills, their sociological background or their motivations are poorly documented: Lam et al. (2011), using users' page gender box and preference setting, for gender studies, report a gender information rate of only 6.5% for the editors of the English Wikipedia. In addition to this, anonymous contributors are, by definition, not registered, when representing more than 90% of the contributions for the French Wikipédia, according to Auray et al. (2007), even if the regular contributors are all registered (ibid). Thus, following Amichai-Hamburger et al. (2008); Yang and Lai (2010); Glott et al. (2010c), to collect a complete set of information on the various participants in the project, we chose to survey these users (and amongst them, the contributors to Wikipedia) regarding their participation in Wikipédia, the French Wikipedia and to link this participation to the socio-demographic variables and the variables describing the first contribution.

About 16,000 people responded to the survey and 13,386 responses were used. Among the people who answered, about two-thirds were non contributors⁵, and just over 12% regular or big contributors (see details in Figure 3.2).

If it is difficult to evaluate the response rate, the number of page view during this period is between 650 and 690 millions, the number of contributors around 65,000 and of active Wikipedians (Wikipedians who contributed 5 times or more in this month), around 5,000⁶. As we captured a bit more than 1,500 answers from regular or big contributors, we estimate that we captured approximately one third of the active Wikipedians on Wikipédia. the very active Wikipedians (Wikipedians who contributed 100 times or more in this month) were around 700 in that period (14% of the active Wikipedians), where, in our sample, the big contributors represent 22% of the regular or big contributors. As previously discussed, these two categorizations cannot overlap, but the same order of magnitude indicates that, if it is not a representative, we have a significant number of very involved contributors. Of course, this also mean that the share of regular or big contributors is bigger in our sample than in the original population, but, as we are focusing on these contributors, this over sampling is rather good for the quality of the results.

⁵We called them "non-contributors" for the clarity of the presentation, even if the use as a reader can be seen as a contribution to the project (Antin and Cheshire, 2010).

⁶stats.wikimedia.org/EN/#comparisons

Table 3.2: Repartition of the surveyed according to their level of contribution to Wikipedia.

Non contributors	Occasional contributors
62.5%	25.0%
Regular contributors	Big contributors
9.7%	2.8%

Our explanatory variables based on the first contribution can be observed only if the Wikipedians have already made a contribution (37.5% of our respondents). This makes possible an over (under) estimation of the dependent variable, due to the fact that some unobserved effects, which have a positive (negative) impact on the probability of having already made a contribution, may have the same impact on the probability of being involved at a more intense level of contribution. To overcome this potential bias, we used a two-stage Heckman procedure which first estimated the probability of being a contributor according to socio-demographic characteristics and a variable representing computer skills in managing complex documents (this variable, which is excluded from the second equation, ensures the identification of the model). Our second step was to estimate an ordered probit model based on the increasing involvement in contribution only considering those who had already made a contribution. The dependent variable is INT_CONTRIB ranging from 1 to 3 with 1 for an occasional contributor, 2 a regular contributor and 3 a big contributor. The convergence to the maximum likelihood in this system of equations can be complicated and computationally demanding. Roodman (2011) proposed a general tool implemented on Stata software which uses GHK algorithm to estimate a full-information maximum likelihood. The procedure models the errors of the two equations (selection equation and ordered probit on contribution) as jointly normally distributed to control for the unobserved effects described above.

Table 3.4, page 48, displays the estimates of this ordered probit model with Heckman correction. Column (6) shows the result of the selection equation and Column (4) estimates our full baseline model. Columns (1) (2) and (3) also use the selection equation in Column (6) but describe the estimate of the ordered probit sequentially introducing information on the form (major or minor contribution), the reason, and the way they made the first contribution. The sequential introduction of additional information concerning the first contribution doesn't change the results in the full model of column (4) (with no significant increase in the standard error of estimators), supporting the assumption that our explicative variables are independent enough.

To ensure that the self-estimation of respondents about their contributions did not bias the results,

Table 3.3: Relationship between the self-evaluation of contribution and the time spent on Wikipedia.

Time spend on Wikipédia (in hour per week)	Share of regular contributors	Share of big contributors
< 1	0.31	0.02
[1 ;5[0.38	0.12
[5 ;10[0.18	0.22
[10 ;20[0.09	0.32
≥ 20	0.04	0.32

we used the time spent on Wikipedia as a robustness check. This information is available for those who have already made at least one contribution. We can see from Table 3.3 that respondents' self-evaluation and time spent on Wikipedia are strongly correlated, as almost 70% of regular contributors declared they spent less than 5 hours per week on Wikipedia while 86% of big contributors declared more than 5 hours. Column (6) describes the result of the ordered probit with Heckman correction for a dependent variable which ranges from 1 to 5, 1 representing less than 1 hour per week and 5 more than 20 hours per week.

The results of this estimate are consistent with those based on the self-estimation of the contribution level.

3.2.2 Results.

Our results confirm, in line with the results found in the literature, that being a male, young or middle age, educated and active strongly increases the probability of being a contributor. Regular contributors, as opposed to others categories of contributors, are male, and not less than 20, but don't differ in the others social aspects. Big contributors are defined by being between 30 and 40 years old, with a master level or more. The negative and significant coefficient associated with the time constraint variable shows that the wikipedians who were not engaged in a professional activity when they started to contribute are slightly more likely to become big contributors (we will come back to this point in the discussion). Not surprisingly, the stronger the involvement in the Wikipedia community the more demanding these characteristics are: time, computer skills, but also experience and knowledge are required to learn how to participate in a collaborative project such as Wikipedia.

The probability of being a regular or a big contributor increases when the first contribution is done one month after having discovered the encyclopedia, this effect remains when contribution is done in the first year. The result is strong and robust to the different specifications of the models, which validates our

Table 3.4: Ordered probit with Heckman selection.

Dependant variable	INT_CONTRIB	INT_CONTRIB	INT_CONTRIB	INT_CONTRIB	HOURS	CONTRIB
	(1)	(2)	(3)	(4)	(5)	(6)
GENDER	0.408*** (0.0706)	0.440*** (0.0679)	0.459*** (0.0655)	0.454*** (0.0706)	0.380*** (0.0697)	0.593*** (0.0258)
AGE16	-0.343*** (0.114)	-0.314*** (0.115)	-0.257** (0.114)	-0.255** (0.118)	-0.320*** (0.113)	0.282*** (0.0602)
AGE 20	-0.266*** (0.0931)	-0.230** (0.0956)	-0.190** (0.0946)	-0.184* (0.0960)	-0.296*** (0.0917)	0.316*** (0.0485)
AGE 30	-0.162** (0.0688)	-0.122* (0.0700)	-0.102 (0.0696)	-0.0893 (0.0709)	-0.162** (0.0680)	0.244*** (0.0395)
AGE 40	0.0380 (0.0729)	0.0732 (0.0719)	0.0849 (0.0710)	0.0964 (0.0743)	0.0262 (0.0729)	0.248*** (0.0466)
AGE 50	0.0739 (0.0774)	0.101 (0.0760)	0.112 (0.0750)	0.103 (0.0802)	0.0785 (0.0785)	0.190*** (0.0516)
AGE+	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
ACTIVITY	-0.0369 (0.0562)	-0.0246 (0.0555)	-0.0290 (0.0553)	-0.0277 (0.0578)	-0.0785 (0.0568)	0.0727** (0.0367)
EDUCATION1	0.0604 (0.0771)	0.0588 (0.0755)	0.0449 (0.0750)	0.0124 (0.0788)	0.0547 (0.0775)	0.111** (0.0436)
EDUCATION2	0.137* (0.0740)	0.130* (0.0723)	0.113 (0.0717)	0.109 (0.0750)	0.0878 (0.0747)	0.127*** (0.0424)
EDUCATION3	0.231*** (0.0812)	0.232*** (0.0795)	0.220*** (0.0788)	0.203** (0.0826)	0.149* (0.0817)	0.264*** (0.0458)
EDUCATION4	0.414*** (0.0856)	0.439*** (0.0832)	0.416*** (0.0825)	0.409*** (0.0865)	0.337*** (0.0854)	0.459*** (0.0437)
EDUCATION5	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
FIRST_CONT1	1.087*** (0.0657)	1.043*** (0.0758)	1.023*** (0.0789)	1.061*** (0.0701)	0.996*** (0.0590)	
FIRST_CONT2	0.470*** (0.0425)	0.445*** (0.0450)	0.435*** (0.0458)	0.448*** (0.0442)	0.435*** (0.0413)	
FIRST_CONT3	Ref.	Ref.	Ref.	Ref.	Ref.	
MINOR_CONTRIB		0.0639 (0.0409)	0.0631 (0.0411)	0.0730* (0.0437)	0.0412 (0.0434)	
MAJOR_CONTRIB		0.275*** (0.0413)	0.275*** (0.0416)	0.280*** (0.0432)	0.264*** (0.0424)	
MOTIV_CURIOSITY			-0.0717* (0.0373)	-0.0943** (0.0397)	-0.0829** (0.0395)	
MOTIV_TEST			-0.218*** (0.0461)	-0.230*** (0.0482)	-0.220*** (0.0476)	
MOTIV_IMPROVE			0.000589 (0.0701)	0.0215 (0.0738)	0.0330 (0.0735)	
HOW_TUTO				0.0608 (0.0394)	0.0518 (0.0392)	
HOW_PEER				0.263*** (0.0768)	0.342*** (0.0756)	
HOW_EASY				-0.0512 (0.0465)	-0.0444 (0.0462)	
YEAR				-0.0175* (0.00919)	-0.0246*** (0.00264)	
COMPLEX_DOC						0.179*** (0.0120)
Constant						-1.980*** (0.0654)
Observations	-11928 13,353	-11904 13,353	-11886 13,353	13,353	-13333 13,353	13,353

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

hypothesis. The marginal effect calculation shows that having contributed the first month increases the probability of becoming a regular and a big contributor by respectively 20 and 21%. Obviously, an early first contribution doesn't explain why a contributor will become at least a regular editor, but this behavior is most certainly a proxy of his motivation and his willingness to get involved in the production of the online encyclopedia.

Considering the nature of the first contribution gives additional information on the motivation and the future status of the wikipedians. The difference between minor and major contribution is both a matter of time spent doing the contribution and complexity of the task performed. An interesting distinction occurs between regular and big contributors. While minor contribution is positively associated with the trajectory of a regular contributor, its become negatively associated with the probability of being a big contributor.

In the same time, a first contribution which is important in terms of investment, like writing or rearranging an article, is positively associated with both regular and major involvement. It seems that those who will become the core editors of the community have a clearly defined purpose since the beginning of their participation and don't "waste" their time with minor improvements on existing article. The standard theory based on the existence of a learning process inside would suggest that the wikipedians start with minor contribution before learning to improve their contribution and participation. At the opposite our result suggests that the biggest contributors started to contribute with a major contribution and didn't need any learning time to get involved in the production of knowledge for Wikipedia.

3.2.3 Discussion.

These results challenge the argument of legitimate peripheral participation. If the trajectory is socially determined and strongly dependent on the first contribution, it raises concerns about the learning process which enables peripheral members to move to the core of the community. Following the title of Panciera et al. (2009), peripheral (readers) and core (contributors) members seems to be born more than made by Wikipedia. This does not mean that there is no learning process, or legitimate peripheral participation.

What Ostrom and Hess pointed out is that online communities decreased boundary between the in and out. It is easier to found and to join this group than a closed, semi hidden, in a word marginalized community of practice such as the witch communities, as showed by Merriam et al. (2003). What we

showed is that, even for joining, i.e. contributing, it seems that the level of effort may vary and is directly linked to the level of the future commitment to the community. This is simple economics: the costlier the first contribution is, the more probable people will keep involving themselves because the understanding of the community, and even the contribution, can be seen as a dedicated asset, hardly valuable outside the Wikipedia context. This is actually one of the basic argument given by Marwell and Oliver (1993) to explain the start of a collective action.

Once the border is crossed, amongst the “in” people, starts a process of leaning, making people evolving from individual contribution and motivation to more collective contributions and motivations. The process of specialization of the editors found by Iba et al. (2010) or Welser et al. (2011), with multipurpose wikipedians, able to participate in the redaction of an article, as the correction of spelling mistake, and wikipedians focused on the editing, may start. To evolve from correcting a mistake to becoming a regular contributor, or an administrator, would be an additional commitment, which would occur for reasons developed during the attendance of the project as the development of this sense of “community”, i.e. the individual acceptance of the rules of the organization, as proved by Pentzold (2011), on his study of the meaning of the term community by the very involved participants of the Wikipedia-1 mailing list (the surveys by Cho et al. (2010a) of 223 English Wikipedians, by Ho et al. (2011) of Chinese Wikipedians, and by Schroer and Hertel (2009) of German ones all prove the link between this “sense of belonging” and the will to contribute; Kittur et al. (2009) showed also how people modify their practices of contributing when integrating the Wikiproject, toward more administrative tasks, according to the group requirement).

This leads to the last question I have addressed I will present in this document: are some communities more efficient to attract such new participants, and to make them contribute. Is it even possible to compare different projects?

3.3 Evaluating Communities of Creation’s Efficiency.

In contrast to studies on open source software (see for instance Crowston et al., 2006; Koch, 2009) that compare projects that use various technologies, programming languages and collaborative tools, the analysis is facilitated in the case of Wikipedia by its structure. This uniformity may help us to better understand, in their difference, what differences are due to process evolution. For each language there is a

separate version of the encyclopedia, a different project, with its own editor community and collection of articles. Importantly, the projects are at different levels of maturity, some quite mature, others still getting started and others somewhere in between. However, they all share the same tool for collaborative edition (MediaWiki) and the same basic rules for collaboration, the “five pillars” of Wikipedia⁷. As well, the global structure of the projects, measured as a network, the nodes being the articles and the links the links between the articles, seems to be about the same, at least for the main projects (Zlatić and Stefanić, 2011).

In a research conducted with Kevin Crowston and Felipe Ortega and accepted in the HICSS 2013 conference Crowston et al. (2013b), we proposed a methodology to evaluate and compare the efficiency with which the different language Wikipedias turn their readers (inputs) into contributors (outputs) and the contributors (inputs) into articles (outputs).

3.3.1 Model and data.

Economists formalize the link between inputs and output as a production function Varian (2005). To be efficient is to reach the maximum possible outputs for a given amount of inputs. In our case, the form of this function is unknown, as are the coefficients relating its components. However, we are not trying to propose a characterization of the Wikipedia production function, but to evaluate if some projects are more (or less) efficient than the others. Since Farrell (1957), this can be done by looking at the “frontier production function”, which describes, for various combination of inputs and outputs, the producers who are efficient, i.e., the ones for which none of the outputs can be increased, without either or several of the inputs increasing or other outputs being reduced, and vice versa.

An additional consideration in analyzing the efficiency of production is the question of “return to scale”, that is, whether a big project may be more efficient because of its size (better known, it is easier to attract new producers, or as explained by Marwell and Oliver, it is the phase where it is rewarding to participate in the collective action).

There are several techniques for estimating this frontier production function. A detailed comparison is out of the scope of this paper, but see Kitchenham (2002) for a discussion of these techniques regarding software production. We choose to use Data envelopment Analysis (DEA) models originally proposed by Charnes et al. (1978), following Koch (2009)’s use in the case of open source software: “these models

⁷<http://en.wikipedia.org/wiki/Wikipedia:Fivepillars>.

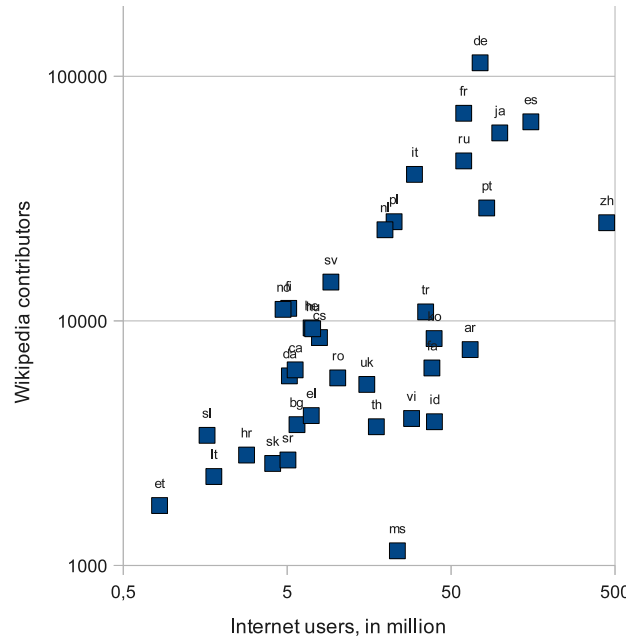
were developed to measure the efficiency of non-profit units, for whose inputs and outputs no clear market prices exist and also no clear evaluation relations” (p. 403). In addition, “DEA can account for economies or diseconomies of scale, and is able to deal with multi-input, multi-output systems in which the factors have different scales” (p. 398). According to the definition of relative efficiency, a DMU (Decision Making Units, here a Wikipedia language project) “is to be rated as fully (100%) efficient on the basis of available evidence if and only if the performances of other DMUs does not show that some of its inputs or outputs can be improved without worsening some of its other inputs or outputs” (Cooper et al., 2011, def. 1.2, chapter 1, p. 3).

As did prior studies of Wikipedia (Lih, 2004; Voss, 2005; Kittur et al., 2007a; Wilkinson and Huberman, 2007; Ortega and Gonzalez Barahona, 2007; Ortega et al., 2009), we relied on Wikipedia internal data to estimate the number of people involved, their characteristics and level of activity. To compute these variables, we obtained the complete database dump with all edits performed in 39 Wikipedias in different languages. These dump files include all required metadata to trace the creation of new articles and individual changes on any page in these Wikipedia projects. In Wikipedia terminology these edits are known as revisions. We were able to retrieve from the dump files the metadata describing each revision, including the identifier of the user who made that edit, its timestamp or the identifier and title of the page that was edited. Data were obtained for each language project for the month of August 2011. For each language we also retrieved an additional file containing information about any special privileges granted to certain Wikipedia users. For instance, in this way we can identify administrators, as well as bots (software programs using Wikipedia accounts to perform routinary or targeted changes in an automated way). The data extraction has been implemented as a software program written in Python to automate this process. This program is part of WikiDAT (Wikipedia Data Analysis Toolkit)⁸. This is a multi-purpose framework that combines Python, R and MySQL with the aim of facilitating Wikipedia data analysis for any of the 280 languages currently available in the free encyclopedia. The use of Python lxml⁹, an efficient library for XML parsing, and multiple sub-processes, let us speed up significantly data retrieval, extracting and computing all metadata and additional information described above (for instance, as far as the English Wikipedia is concerned, 444,946,704 revisions in 27,023,430 pages were analyzed in approximately 44 hours). This massive data analysis allowed us to develop more precise data than those

⁸<http://libresoft.es/node/564>

⁹<http://lxml.de/>

Figure 3.1: Number of contributors versus Internet population.



presented by the Wikimedia Foundation¹⁰ as far as the edits and the contributors are concerned, and to include new, original data, the number of FAs and the number of new FAs by month.

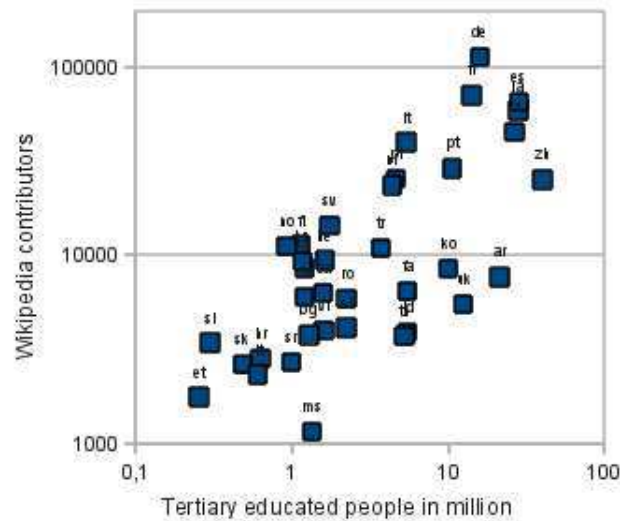
3.3.2 Results.

Plotting the data shows a strong (but not perfect) correlation between the total number of Wikipedia contributors and the Internet population (figure 3.1), and the total tertiary-educated people (figure 3.2). Using the DEA model, we are able to determine the different levels of efficiency in the conversion of these inputs to the Wikipedia community of contributors.

The results for this analysis are shown in Figure 3.3. The projects are listed in decreasing order of size). The bars indicate relative efficiency. The longest bars, representing 100% efficiency, are for projects that are on the efficient frontier, creating the most outputs from the particular mix of inputs. Shorter bars represent projects that use a similar mix of inputs but produce less outputs than other projects. The results indicate varying levels of efficiency in converting potential editors to actual editors. Specifically, language projects such as Malaysian (ms), Arabic (ar) and Chinese (zh) have many fewer editors than would be

¹⁰<http://stats.wikimedia.org/EN//>

Figure 3.2: Number of contributors versus population with a tertiary education.



suggested by the population of Internet users who could become editors, while Estonian (et), Hungarian (hu), Norsk (no) and Finnish (fi) show high efficiency in recruiting editors.

As far as the return to scale is concerned, Table 3.5 presents the sign of the return to scale variable, v_0 . It seems that the biggest efficient projects have entered in a decreasing return to scale phase ($v_0 < 0$), suggesting increasing difficulty to recruit new Wikipedians. In the other hand, the smaller projects, when they are efficient, seem to be still in an increasing return to scale phase.

The second model examined the production of edits to articles, of new articles and of new articles and new FAs. The results are shown in figure 3.4 (production of edits, articles and FAs). Yellow bars show the efficiency of producing edits, navy blue bars, efficiency in producing new articles, and red bars, efficiency in producing articles and new FAs.

The difficulties of the main projects to maintain a constant level of activity as the stock of articles increases appears clearly, as the return to scale is systematically negative for the larger projects (see Table 3.6). But beside the Japanese project, the main projects are still efficient in terms of level of activity. On the other hand, projects that apparently find it difficult to recruit editors may still be efficient in the converting the effort of the workforce available into edits and articles, as is the case for the Malaysian (ms) and Farsi (fa) language Wikipedias.

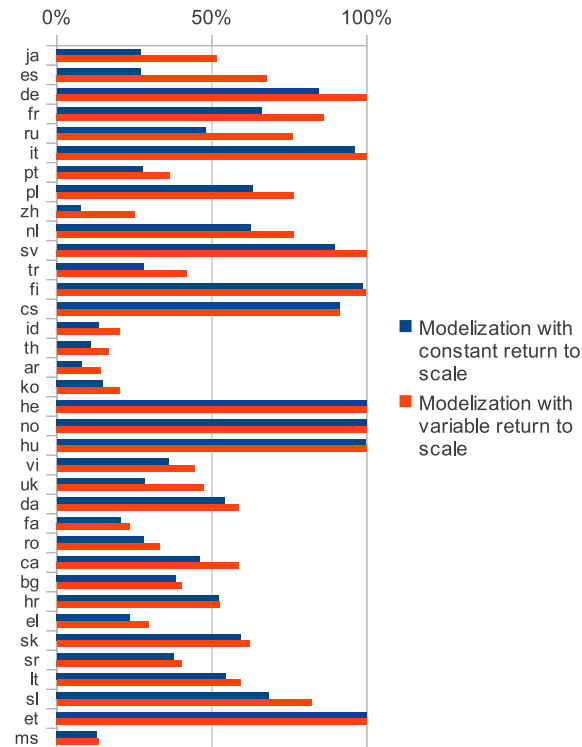
Figure 3.4 helps to explain if edit activity is due to a high level of production of articles or the results of activities that consume activity but do not lead to an increase of the stock of article, either positive,

Table 3.5: Return to scale for the recruitment of contributors.

Japanese	ja	-1.57
Spanish	es	-1.60
German	de	-0.04
French	fr	-0.11
Russian	ru	-0.12
Italian	it	-0.12
Portuguese	pt	-0.17
Polish	pl	-0.14
Chinese	zh	-0.15
Dutch	nl	-0.10
Swedish	sv	-0.45
Turkish	tr	-0.29
Finnish	fi	-0.03
Czech	cs	-2.19
Indonesian	id	-0.65
Thai	th	-0.38
Arabic	ar	-0.73
Korean	ko	-0.09
Hebrew	he	-0.08
Norwegian	no	0.02
Hungarian	hu	-0.14
Vietnamese	vi	-0.36
Ukrainian	uk	-0.64
Danish	da	-1.31
Farsi	fa	-0.62
Romanian	ro	-0.15
Catalan	ca	-0.78
Bulgarian	bg	-0.49
Croatian	hr	0.34
Greek	el	-0.75
Slovak	sk	0.42
Serbian	sr	-0.16
Lithuanian	lt	0.13
Slovenian	sl	0.15
Estonian	et	-0.19
Malaysian	ms	-0.56

In red when the project is efficient.

Figure 3.3: Efficiency in recruitment of contributors.



Note: projects are listed in decreasing order of size.

such as a focus on improving existing articles, or negative, such as bureaucratic discussion or even edit wars. For instance, the high level of edits in the French (fr) and the German (de) language projects seems to be due to a focus on FA production rather than the production of new articles, for which those projects seem rather inefficient (even after having taken into account the decreasing return to scale in the model). On the other hand, projects of intermediate size, such as the Russian (ru) or the Italian (it) Wikipedia, are still very active on the level of new article production.

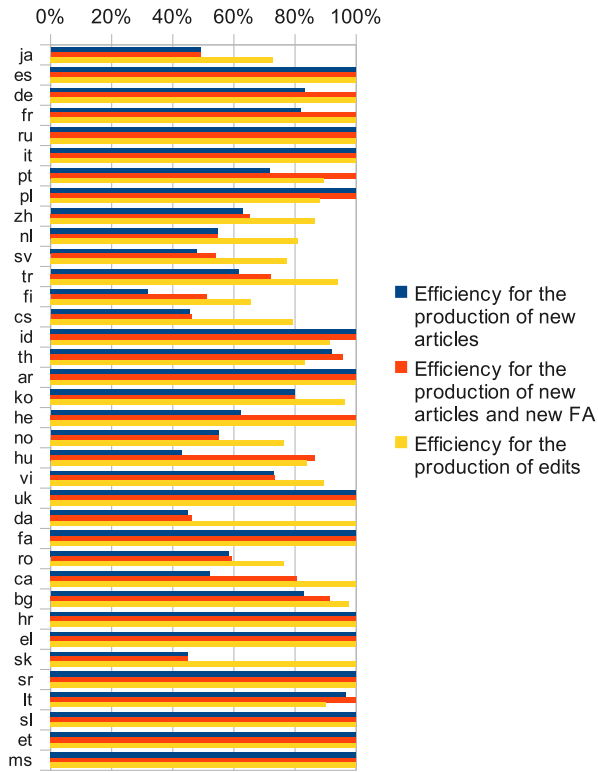
Finally, the level of edits seems to be a good indicator of the level of final production, as few of the projects which are inefficient at the edit level are efficient at the article or article and FA level. The exception to this rule are the Lithuanian (lt), Portuguese (pt), Polish (pl) and Indonesian (id) projects. A possible explanation, given by van Dijk (2009), and proven for the Indonesian project by Soekarno and Giri (2005), is that an important part of the articles in those projects are directly translated from English, and these articles require less editing to be published.

Table 3.6: Return to scale for the activities.

Projects	Number of edits	Production of new articles	Production of new articles and of new FA
ja	-0.02	-0.95	-0.95
es	-0.01	-0.36	-0.36
de	-0.99	-1.20	-0.93
fr	-0.49	-1.22	-0.98
ru	-0.05	-1.00	-0.12
it	-0.17	-0.57	-0.47
pt	-0.10	-0.80	-0.89
pl	-0.10	-0.67	-0.71
zh	-0.03	-0.40	-0.33
nl	-0.13	-1.01	-1.01
sv	-0.22	-1.43	-1.26
tr	-0.25	0.13	-0.82
fi	-0.10	0.34	0.10
cs	0.10	0.69	0.25
id	-0.05	0.21	0.42
th	0.81	1.43	1.32
ar	-0.36	0.00	-0.59
ko	0.02	0.32	0.32
he	-0.09	0.58	0.12
no	-0.21	0.15	0.15
hu	-0.03	0.04	0.10
vi	-0.06	0.19	0.20
uk	-0.01	0.15	-0.34
da	-0.10	1.76	1.63
fa	0.01	0.21	0.06
ro	-0.22	0.28	0.35
ca	-0.18	0.22	-0.26
bg	0.20	2.76	1.22
hr	2.77	3.23	2.77
el	0.26	2.12	1.34
sk	0.19	2.61	2.61
sr	-0.56	0.53	0.45
lt	0.49	0.80	0.22
sl	-0.10	3.75	0.94
et	0.77	1.56	1.90
ms	0.65	1.40	2.03

In red when the project is efficient.

Figure 3.4: Efficiency for the production of edits and articles (new articles and new FAs).



Note: projects are listed in decreasing order of size.

3.3.3 Discussion

Our analysis showed striking differences in efficiency in the two processes among the projects. For the differences in efficiency in recruiting participants, the size of the project seems to matter, as all the larger projects are assessed as being inefficient. In the model adding a factor for return to scale, the larger projects increase their performance, with a negative return to scale (V_o being negative). In other words, it may simply be that the largest projects have reached a size where it is harder to make a new contribution and so harder to recruit new Wikipedians.

Nevertheless, there remain striking differences in efficiency among the smaller projects. We proposed two possible explanations for these differences. First, many of the less efficient projects have a lower level of tertiary-educated people compared to the efficient group. This difference could be a key to explaining

the low efficiency of recruitment. A second hypothesis is on the control of the information: many of the low-efficiency projects are tied to countries where the Internet and the production of information is more closely controlled by the authorities than in the efficient group. It may be that freedom of expression is pre-requisite for efficient recruitment of editors. Zhang and Zhu (2011)'s recent study on the Chinese Wikipedia gives arguments for this hypothesis.

As for production, it appears that some of the difference can be attributed to the level of maturity of the projects. Newer projects have fewer articles and so it is easier for contributors to find topics that have not been covered. For the larger and older projects, is the gap between efficiency in editing and in creating new articles because work is being directed to improving the quality of the article, or is it a sign of inefficiency (ineffective edits)? The current evidence is inconclusive. Lih (2004); Wilkinson and Huberman (2007), confirmed by Ortega (2009), found that after taking into account age and visibility (using Pagerank as a proxy), FA status could be predicted by an increased number of edits or number of editors. Kittur et al. (2007b) found that in 2001, 90% of edits were done in the main namespace on the English Wikipedia but that this number dropped to 70% by June 2006, suggesting that the efforts of editors are being diverted to less productive activities. However, Wilkinson and Huberman (2007) found that articles with more discussion on their Talk page were generally ranked higher in quality according article ratings, suggesting the tradeoff between simple production and efforts to improve quality.

The work presented here lays the groundwork for additional research. First, future work should include outputs along additional dimensions, considering factors such as article size and quality, as well as the whole organization of the encyclopedia, which are the usual dimensions for analyzing documents in library studies (see for instance Rector, 2008 on a comparison of Wikipedia with other encyclopedias). Second, this analysis would benefit from distinguishing more finely among different kinds of editors, beyond the three levels of very active, active and other used here. On the flip side, the analysis should also consider the contributions of non-registered editors. Anthony et al. (2009), examined contributions from registered and non-registered users, showing the importance of anonymous contributors in the total production. Finally, moving down from the level of an entire language project, this kind of analysis might also be done at portal or subject level. Poderi (2009) does a similar analysis though only on a small subset of article. They obtain counter-intuitive results, as it seems from their analysis that the subject having the more feature articles (high-density subjects in his terminology) have longer articles, but fewer edits and

contributors than the low-density subjects, while the ratio between major and minor edits is the same in the two groups. It seems also that there is more often a single major editor in the high-density subject articles. This result have been confirmed by Kittur and Kraut (2008) who found a positive impact on an article's quality from an increase in the size of the number of editors only when a small core of editors performed the majority of editorial work.

This discussion opens the way for the last part of our document, the research paths opened by the work done, and more precisely the question of the relationship between firms and communities.

Chapter 4

Future Research: the Market and Communities of Creation: a Work in Progress.

As presented in the introduction, in my future research I want to deeper my work started on the link between market and online communities. This means better understanding how each piece of the jigsaw fits: better understanding open-source business models requires a better understanding of the adoption of open-source products; better understanding firms' participation in online communities requires better understand the reasons why they participate, and knowing that they participate via their employees, which is what both employees and employers expect from this participation; finally some work can be done to better understand how these communities work, regarding the cooperative production of the pieces of knowledge, and regarding the competences developed by the participants (in other words, the two things people are seeking in those communities: pieces of knowledge and social/cultural capital).

4.1 Open-Source Demand.

4.1.1 (Information) Technology adoption.

Theorists have proposed two general answers to the question of the diffusion (i.e., more widespread adoption) of technological innovations, with few interaction between them (Dedrick and West, 2004). This situation is unfortunate, as, as Fichman and Kemerer (1993) showed, their criteria can be combined to explain the domination of a technology. On the one hand, Roger's model of innovation adoption (1983), ap-

plied by Fichman (1992) in the case of information technology, focuses on individual decisions based on inherent characteristics of the innovations. As Fichman states, “Innovations possess certain characteristics (i.e., relative advantage, compatibility, complexity, trialability and observability) which, as perceived by adopters, determine the ultimate rate and pattern of adoption” Fichman (1992, p. 7). On the other hand, economists have focused on the increasing returns to adoption that lead to standardization (Katz and Shapiro, 1986b, 1994) as the explanation for increasingly widespread use of a particular technology within a market. The reasons for these increasing returns are various (Arthur, 1989b,a): decreasing costs for adoption (due to economy of scale in production, learning effects from both producer’s and adopter’s sides), or increasing perceived value of the adoption (due to learning effects, too, but also to adoption by others that produce network externalities and technological interrelations). In this view, a technology is adopted ultimately because it has become cheaper than the alternatives. A further complication is that technologies are frequently “re-invented” over their diffusion as current and new players offer competing products. Swanson’s studied the diffusion of an innovation (1994), and noted that “an IS innovation is likely to evolve such that it is increasingly tailored (or even transformed) by means of new features which accommodate the adoption of newer, related innovations”. In other words, the diffusion process is often a double diffusion, first of a dominant design, and second of the various specific technologies offered. Sometimes, competitors propose different designs and there is a standard competition to determine which will become the dominant design (e.g., VHS vs Betamax); sometimes there is a normalization process which settles the dominant design and competition is then between solution producers (e.g., GSM). As pointed out by Rogers (1983) or March (1981), it may be a new entrant who is the one to impose the standard (Geroski, 1995; Markides and Geroski, 2005), indeed, more surely that the first to propose the technology.

Is the diffusion of open-source solutions due to a better re-investment of the benefits of adoption, as we defended it a decade ago (Dalle and Jullien, 2003), allowing open-source products to better address new adopters during the diffusion process, and thus to a better way to create the standard?

The arguments for this thesis is that the open source organization presents a very efficient solution to the specific incremental process of innovation in the software industry, thanks its modular organization (Bessen, 2005). **Or is the diffusion simply due to the fact that, as a technology has become mature, a cheaper, free license-fees solution imitating the standard offer can be developed, in a market where**

process and production innovation is harder to protect? We will take ERP's¹ example to develop the kind of studies that could be conducted in the future regarding OSS adoption by companies.

ERP market seems to be a good field of investigation, knowing that these two arguments are advanced to defend the adoption of open-source ERP by SMEs, in two theoretical works. ERP is a very important IT for companies today, even for medium-sized companies, but these systems are also very expensive.

4.1.2 Open-source ERP as a case study for open-source adoption.

An OSS solution should help to facilitate adoption, because of the decreasing cost of adoption, according to Johansson and Sudzina (2008), but also because of “increased adaptability”, and “decreased reliance on a single supplier”, according to Serrano and Sarriei (2006). In addition to this, ERP seems to be a technology that firms, and especially SMEs want to try before adoption (Ramdani et al., 2009), which should favor open-source solutions.

I am not going to detail here all the variables mattering in the explanation of the adoption of a technology, and I will rely on Crowston et al. (2013a). Fichman and Kemerer (1993) proposed a framework to study the variables impacting the adoption of a technology by a company, used by Dedrick and West (2004), which consists in regrouping the impacting variables into three main categories, the technological factors, the organizational factors, and the environmental factors. As these last authors stress, it is more a matter of classification than a stable framework, and we will use it as so. Regarding technological factors, and relying on Rogers, in a meta-analysis of prior studies, Tornatsky and Klein (1982) concluded that three of variables were consistently linked to technology adoption: compatibility, relative advantage, and complexity. However, regarding the open-source solutions, Dedrick and West (2004) concluded that trialability also matters. Regarding environmental factors, Robertson and Gatignon (1986) noted that a variety of competitive effects in the technology consumer's industry (competitive intensity, demand uncertainty, professionalism, cosmopolitanism) and within the technology supplier's industry (level of competitiveness, reputation, R&D allocation, technology standardization) affect the diffusion. But, as pointed out by Haddara and Zach's review of the literature on ERP in SMEs (2011), and by Morgan and Finnegan's and Dedrick and West's ones on open-source (rep. 2007 and 2004), the organizational factors

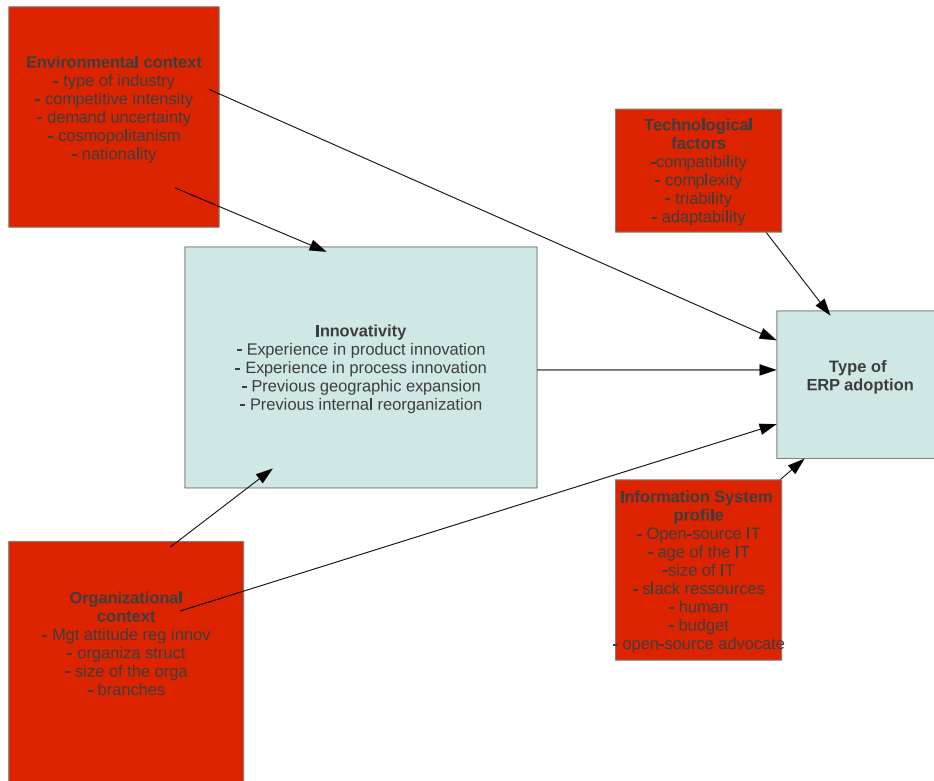
¹For enterprise resource planning. Rosemann (1999) defines an ERP system as a customizable, standard application software which includes integrated business solutions for the core processes (e.g. production planning and control, warehouse management) and the main administrative functions (e.g. accounting, human resource management) of an enterprise. In the following, we will rely on this definition, but also on the products which define themselves as such, proprietary, such as SAP and SAGE vs open source, such as Compiere, ERP5, Open Bravo.

seem to be preponderant in the adoption of such innovations, and we will not detail the environmental factors here.

Other variables explaining IT adoption have been proposed, such as the management attitude regarding innovation and the structure of the organization, but with weak results (Haddara and Zach, 2011), even if Chang et al. (2010) found positive correlation between CEO attitude towards IT adoption and CEO's IT knowledge to have significantly more chance to adopt ERP. However, this does not close the list of explanatory variables for the adoption of an open-source ERP, as if the link between IT investments and productivity is today well established (Dedrick et al., 2003), Eric Brynjolfsson pointed in an article in the Technology review (2004), to summarize his research on the link between productivity and digitalization (Brynjolfsson and Hitt, 2003), that "The unsung heroes of the IT revolution have not been the microchip and the Web browser, but rather the creative, diligent, and painstaking work done by those who have been rethinking supply chains, customer service, incentive systems, product lines, and 1,001 other processes and practices affected by computers. Investments of intangible capital constitute the real source of today's productivity growth".

The hypothesis is that two types of firms and of reasons for firms to invest open-sources ERP solutions: firms or organizations for which the adoption of the ERP can be seen as the **adoption of an innovation by followers** ("ERP software has become a commodity", Stratman, 2007), and thus which are mainly concerned by the cost of the solution, and second organizations which seek for **new competitive advantages as first movers toward more adaptable and business adapted solutions**, thanks to the modularity of the open-source solutions. Actually, the trialability should be considered by both adopters, but not for the same reasons. For the first group, it should be a way to reduce the risks due to the uncertainty of the solution, when, for the second, it should more linked to tests to choose the best solution. The fact that this adoption follows or is coordinated with process innovation should favor the 2nd kind of adoption of ERP. **The first type of adopters should be of smaller size, in less competitive environment, more centralized. The second type should be more innovative firms, having the slack resources to try the open source offers, but also a more committed management.** Control variable are Robertson and Gatignon's ones (1986), competitive intensity, demand uncertainty, professionalism, cosmopolitanism within the organization's industry and within the technology supplier's industry (level of competitiveness, reputation, R&D allocation, technology standardization), and cultural (country-level) factors, such as uncertainty avoidance, power distance individualism (based on Hofstede, 1991), IT com-

Figure 4.1: Open-source ERP adoption framework.



In red, the explanatory variables, in blue the variable(s) to be explained.

petence, economic development, based on Qu et al. (2011a).

The general framework of the proposed research is summarized in Figure 4.1.

4.1.3 Data collection.

I plan to answer these research questions by surveying evaluators and adopters of an open-source ERP at international level (France and Europe, the USA, meaning that the questionnaire will be translated in English and French). The survey instrument will be developed to operationalize the concepts discussed above, in addition to demographic variables. The construction of the survey will rely on several sources. In addition to two questionnaires studying the adoption of ERP: Buonanno et al. (2005), already mentioned for the description of the company, and Deltour and Trémenbert (2010) on the internal IT capacities (people slack resources), the key element for the competitiveness of the firm, the ERP mod-

ules implemented, the competitive objective pursued while implementing the ERP, global investment in IT, other organizational innovations made in the same time. I will rely on Robertson and Gatignon (1986) to complete the control variables, and on Qu et al. (2011a) to express the cultural (country-level) factors. I will use Damanpour (1991) to check if the two proposal questionnaires take into account the whole set of variables explaining organizational innovation, and will complete with his proposals, if it is not the case. The survey will be designed to be administered via an online questionnaire. The level of analysis of the survey will be the organization, but I will rely on a key informant to report on behalf of their organization. The population of interest is those who have considered and perhaps adopted an OSS ERP system. To reach this population, I plan several strategies. First, I plan to work with companies that supply different OSS ERP systems like Phidias (Phidias is European partner of the year for the open-source ERP Open-Bravo, the world's leading web-based Open Source ERP solution), or Open-Bravo.com or ERP5 provider Nexedi, as ERP5 is one of the most studied solution in the academic area (Carvalho, 2009), and the Compiere.com company. This would allow me to survey users and potential users of the three main open-source ERP. Second, surveys through companies will be complemented by a request in the LinkedIn group lists, which are quite active (Compiere Community Group - Open Source ERP / CRM, Enterprise Open Source Community, 691 members, Openbravo, 888 members). I will control also the answer addressing other ERP groups on LinkedIn, such as ERP community (29,000 members). For comparison, I can also survey members of groups dedicated to proprietary ERP software, such as the SAP community group (137,000 members), Sage ERP Solutions and Oracle ERP User Network (two closed groups).

Of course, the result of this study will have to be extended to other kind of software, to have a broader vision of the link between the impact of the software adopted on company's organization and reason why choosing open-source solution. Regarding the importance of the factors (price or flexibility), open-source solution providers may not be able to have the same business strategy, and thus, the same involvement in open-source communities. This involvement, its level, its measurement, and its impact on the management of the employees participating in open online communities is the second path of research I want to discuss here.

4.2 FLOSS Business & Participation Strategies and Employee Management.

Coming back to open, online communities, the fact is that more and more companies contribute to free software, or more accurately, more free software contributors are paid to do so (Lakhani and Wolf, 2005; Vicente, 2008). **Why and how do they allow their employees to participate in those communities? Is it because of business purpose, as I argue, or because this allow them to have better, more committed employees?**

A little detour via Corporate Social Responsibility (CSR) studies may make my point: do open source/CSR companies get involved in open source/CSR because communities and consumers ask for this, as shown by Maignan and Ralston (2002)? Is it a niche strategy for some companies targeting specific customers interested in open source software properties as developed in this section, or/and an evolution companies have to include in the definition of their business strategy, as advocated by Porter and Kramer (2006), still looking at CSR. In that perspective, Waldman et al. (2006), via an extensive cross-national study of 561 firms based in 15 countries, on five continents, showed that beyond corporate strategy, cultural values impact the way top management integrate CSR values in their strategic thinking (see Aguilera et al., 2006, for a review of the literature on comparative studies on the topic). Employees' perception of the importance of CSR matter too, and can be used by the management to develop a belonging to the firm perception, for instance (Rupp et al., 2006). In return, the employees, as other stakeholders, may advocate considering CSR in firms' business decision, with the result that, as for open-source adoption, CSR decision are multi-level (see Aguilera et al., 2007 for a theoretical framework to analyze this impact).

The question is, from my point of view, threefold: **better defining firms' business strategy and its impact in terms of participation in the community** (Fitzgerald, 2006); **employees' work strategy and its impact in terms pf participation in community**; **how they negotiate and align their potential orthogonal goals** (on that topic, see the analysis of the cultural differences between proprietary developer and open-source developer by Rolandsson et al., 2011). I will present those three points before presenting the methodology I propose (data collection and analysis) to address them.

4.2.1 FLOSS Firms' strategies.

FLOSS companies choose what “they want to disclose, to open source” (Henkel, 2006b). If in 10 year-old study (Jullien, 2003), I partially validated the model we proposed with Jean-Benoît Zimmermann, **the evolution of the computer market plead for an updating of our findings, and for a closer evaluation of the firms' open-source investment.** At first glance, the software as a service (SaaS) trend, because it disconnect even more than before buyers' investment from their technological choice may make the choice of an open-source solution easier for the provider, which is competing on (3A) services and price. In the other hand, it may make it harder for these solution providers to externalize their development to VH users, as these users would not install the service. Some open-source actors, such as ExO², seem to develop a two-channel offer, maybe to maintain this link, in addition to the externalization of their core developments to low-cost countries. **Studying firms' business strategy and their involvement in FLOSS communities remains a prerequisite before going into the analysis of the relation between firms, employees and communities.**

4.2.2 Employees' FLOSS strategy.

As far as corporate perspective is concerned, and as explained before, the open-source firm employees may be seen as representing and acting for their employer when participating in the development of such communities.

The “Career” becomes double for open-source developers: they must simultaneously manage the evolution of their participation in the community, and their professional career (Vicente, 2008). I showed (Demazière et al., 2005; Jullien et al., 2011) that if the signaling perspective (the credit, discussed in chapter 3) is not pertinent to explain the initial involvement in open, online production, it becomes important for the main involved. The professional constraint may have an impact on developers commitment and motivation. Can we, for example, still speak of freedom (of expression, choice, contribution), which is “one of the foundations of the free software culture” (Scacchi, 2007) and a well-known incentive to participate, where individuals are paid for relaying, putting into practice the strategy of their employer? **Corporate behavior may contradict the value of the community and thus have an impact on their effectiveness as ideology has an influence on team effectiveness of free software** (Stewart and Go-

²<http://www.exoplatform.com/company/fr/home>

sain, 2006a). In summary, free software employees are likely to encounter a contradiction, a “tension” (in the sense of Thévenot, 2007) between their commitments vis-à-vis the community and vis-à-vis their employer.

How people deal with this tension is mainly the blind spot of today’s research, apart from Rolands-son et al.’s analysis of the difficulties for employees to switch from a proprietary development model to a more open one (2011).

4.2.3 FLOSS firm’s management strategies.

The specifics of this situation is mainly related to three actors: a developer at the same time employee of a company and also involved more or less strongly in a open-source development activity, his/her hierarchy which must ensure the achievement of organizational goals and thus distribute the right incentive to him/her, and finally, a community of peer supporting the development of free software. But **is the control of the asset “community” the unique reason why these companies let their employees get involved?**

Other companies such as Google give some free time to pursue personal projects and, amongst these projects, and **personnel economics theory proposes other reason for firms to let their employees having personal occupation during their work-time**: compensation, self training, competence signaling, amongst other (see Lazear and Oyer, 2013 for a review of the literature on the domain). This echoes the managerial problems posed by the transformation of closed business in open company and more generally the management of innovation in companies. Management of human resources, the peculiarity of the study of innovative teams and the importance of interactions to foster this innovation was highlighted by Defélix et al. (2005), which also offers a literature review on the subject. This was also emphasized in a field related to the production of videogames (Parmentier and Picq, 2006).

Personnel economics proposes various explanations for the fact that firms would favor such an involvement, in addition to strategic reasons already evoked, and which can help mitigate or extend the simple strategic involvement: compensations to salary have to be explored (salary over benefit trade off, see Lazear and Oyer, 2013, pp 30-31, for a way to measure it), but also recruiting process where this compensation is used to attract the best developers, in this innovative environment, something which seems important in the software sector (Andersson et al., 2009). In that sense, and following Marino and Zábojnák (2008), community involvement would be seen as a perk, closely linked to the worker’s job, which

would improve her motivation. This can also be viewed as a kind of firm sponsored personal training, which at least for non-skilled workers, is proved to be good for the company, which can appropriate a part of the rent created by this training (Acemoglu and Pischke, 1998), and makes it harder for the training person to leave the company during the training period (Flaherty Manchester, 2008). However, the same models prove the appropriation and retention to be more efficient when the quality of the employee remains the private information of the employers. When participating in an open community, the participants may reveal their quality to the competitor, resulting in an decrease of investment in training from the firm (Acemoglu and Pischke, 1998), as open source communities make the market for high skilled developers more liquid. At the same time, the time spent by the employee remains the firm's private information, making the evaluation of an employee's productivity harder for the competitor.

These hypotheses are to be tested both survey the employees and the management of the companies.

4.2.4 Model and data collection: the question of the measure.

It is, actually, not that easy to evaluate the participation of firms in FLOSS projects. They do not contribute by themselves to the projects, but via people hired to do so. As already said, some works exist on that topic and they start from firms' people's behavior in a community, so using the data produced by the community. This suffers from some limitations. The way scholars collect data on that topic is to look at the data produced by the communities (email lists, source code produced and signed, etc.) People are said to belong to a company when they sign with a company address. The main limitation of this is the exhaustively and the coherence of the data collected: some developers do not contribute using their employer's address, and others do but without their employer's agreement. However, when firms commit themselves to a community (as IBM could have done to the Linux community), this technique may be usable to track their involvement and their evolution in time. And, as Gleave et al. (2009) pointed out for people, looking at what firms do does not provide information on why they do it, or if this is driven by the management, in a word, of firms' motivations.

To test the weight of each explanation of people's motivations for participating can be done via a qualitative and quantitative mixed approach, such as the one Shah (2006) did: after case studies in open-source companies such as XWiki, Mandriva, or Linagora, a survey of the in-firm open-source developers and of their management would be conducted. According to the analysis presented here,

regarding firm questionnaire a point that should particularly be studied, in addition to the reasons for the free time given by firms to their employees, if any, is the **missions given by the firms to developers participating to a community**: are they hired to participate, are they hired because they participate, and do they have a specific goal (to reach a certain level, to take charge of a certain part of the project), or is participation the only thing required? Regarding the employees, personnel economics theories suggest that evaluating the value of the free time as a perk, i.e. measuring the social, cultural capital employees gain by participating in those communities may be a breakthrough.

As pointed out by Lazear and Oyer (2013), the complementarity and redundancy between the different dimensions of community involvement as a perk make it very difficult to measure this perk. **If both models are less developed in my mind than the adoption model, both firms' strategies regarding FLOSS involvement and employees' participation in these communities are multi-factors, and the parallel with CSR can be fruitful as its theories integrate a multi-stakeholder perspective and propose models to study them.** In addition, complementarity and redundancy between the variables suggest that traditional econometrics measurement may not be appropriate. **Meyer and Ponthière (2011) proposed innovative methods and measurements in that case, where people are asked to rank multiattribute hypothetical combinations of factors and Choquet integral-based multiattribute value theory is used to elicitate the ranking of preference.** Having Patrick Meyer as a colleague at Télécom Bretagne will facilitate the transfer of such an innovative methodology, in between experimental economics and traditional questionnaires. This is, to my mind, one of the assets I have in developing research in economics in a school of engineering, a point I will develop in the conclusion of my work.

4.3 Efficient Community, Good Contributor.

As already said in the introduction, the flip side of the coin is the online community. Companies have to evaluate the projects before investing in, and community managers may also want to monitor their functioning³. Finally, studying these community may help us to better understand how groups work, how (virtual) collaboration and leadership can succeed, something of growing importance for firms (Crowston et al., 2010; Hernandez, 2012).

³See the evaluation of FLOSS solution by CIGREF, http://www.cigref.fr/cigref_publications/RapportsContainer/Parus2011/Maturite_et_Gouvernance_de_l_Open_Source_CIGREF_2011.pdf, and more generally, the Open-source software assessment methodologies developed, most of the time, by service companies, on http://en.wikipedia.org/wiki/Open-source_software_assessment_methodologies.

4.3.1 Measuring efficiency.

Behind efficiency, one may put several things, though. For instance, the process(es) of collective activity can be analyzed through several aspects: e.g. group composition, degree of collaboration, coordination, roles distributions, dynamics of the collective process, interpersonal processes. The outcomes can be apprehended at several levels: a productive level (various characteristics of constructed epistemic knowledge such as completeness, creativity... as well as its utility/usability, more user-oriented), a collective level (e.g. team building, construction of rules and collective norms...), and a developmental level (learning and development of individuals). The relationship between process(es) and product(s) can be approached in terms of efficiency but also with the more complex integrative concept of “quality” to understand what links can be made between “good” process(es) with product(s) of “quality” (see the global scheme presented in the introduction, figure 1.1, page 11). To stay on track with my perspective (link between the communities of creation and the digitalized economy), I will focus **on perspectives on the evaluation of these communities (efficient/ non efficient) and on the career of the participants**: what do they learn in those communities, what are the ways to measure the (cognitive, social) capital they earn in contributing?

Regarding the efficiency of the production, Stefan Koch proposed several studies on the measure of the efficiency of FLOSS communities, and defines “efficiency” the capacity of turning contributors into lines of codes, but also into bug fixing, for instance (Koch, 2008b, b). As this list indicates, something he pointed out, the outcomes of a community have to be measured according to several dimensions, without a clear view of the importance of each factor, leading him to propose to use Data Envelopment Analysis techniques to estimate which projects are “efficient” without assuming a form for the production function. As already said Kevin Crowston, Felipe Ortega and I transferred this technique to Wikipedia studies. Regarding this work, improvements have to be made, including outputs along additional dimensions, considering factors such as article size and quality, as well as the whole organization of the encyclopedia, which are the usual dimensions for analyzing documents in library studies (see for instance Rector (2008) on a comparison of Wikipedia with other encyclopedias). This will require a better measure of article quality⁴.

⁴The most comprehensive attempt to develop criteria to judge article quality on Wikipedia may be the ones by Stvilia et al. (2008) and Lewandowski and Spree (2011). Stvilia et al. (2008) looked at the information quality process both in the organization (number of editors, of edits, ratio between edits in talk pages and in content pages, etc.), and in people’s interaction (via a content analysis of a set of feature articles’ talk pages). Lewandowski and Spree (2011) extend these criteria to 13 criteria (see p. 126 of their

Beyond this natural, but quite specific extension of my research, I think two points have to be investigated more deeply. As these two points are connected in the techniques used to extract information, and because they are interdependent in their results, I think they have to be treated together: **what is a good virtual team, in term of size and of people know-how? And what knowledge and connexion people develop in participating in those teams, how this evolve over time.**

4.3.2 Looking for the “good” virtual team.

Regarding the first question, Stefan Koch⁵ also tried to estimate the effort participants put into the production of FLOSS using classical production models and stressed the fact that FLOSS organization seem to require less contributors than a private model, because of a better organization (people participate in what they are interested in, no over-staffed management), and because of a significant, albeit unmeasured contribution from users (bug report, documentation writing, etc., Koch, 2008a, a). One of the main results found is that modularity is a key factor of success in open-source projects, as it allows a very little team to work on the same file, avoiding the classical congestion problem in computer programming (at a certain level, adding more people to a project may only slow it down more, but when there are a lot of files to work on, they can dispatch accordingly). This result has been somewhat confirmed by Kittur and Kraut (2008) in the case of Wikipedia, as they found a positive impact on an article’s quality from an increase in the size of the number of editors only when a small core of editors performed the majority of editorial work. Ransbotham and Kane (2011) went further, still on Wikipedia, showing that a fine tune of experimented editors and fresh newcomers increases the likelihood for an article to become a feature article.

In the context of creative industries and workers (Broadway musicals and university research teams), Uzzi and Spiro (2005); Uzzi (2008) **proved that for a creative group to be successful, it needs to fine tune the level of newcomers, for fresh ideas, in an already constituted group (for trust and common sharing, or “cohesion”)**. They show that there is what they call a “Q”-level, “bliss point”

work for the complete list), drawn from data analysis (length of the article, existence of references, etc.) but also human (expert) evaluation of the quality. They show a correlation between these criteria and the rank in search engine, with a good correlation but a strong dispersion. In both cases, the automation of the methodology to a whole project, not to say to different languages, seems impossible. There are efforts to automatically analyze the articles, but these are currently not yet enough effective to be of use (Fong and Biuk-Aghai, 2010). Indeed, even the fact that an article is a FA (feature article) is not coded in the projects’ data base, and instead has to be extracted revision by revision from the text of each article in the projects’ dump.

⁵ A presentation of his finding, and a review of the literature on FLOSS organization and efficiency can be found in Koch (2011).

in the ratio between those newcomers and experimented people (more is too much, less is too few) for these creative teams to be the most successful. Interestingly, **this seems to confirm also certain results of personnel economics studies on team working, which show that heterogeneous teams are more productive than heterogeneous isolated workers, in the case of low level skill workers** (Mas and Moretti, 2009; Hamilton et al., 2003). But their main argument is quite different, as they argue that social pressure makes the less productive work harder, even if Hamilton et al. stressed that the internal learning process may also play a role. In a condition where social pressure and help are explicit in the discussion list, the reasons for this increase in productivity may be easier to unravel. This argument is closed to what I defended with Sylvain Delean: close social monitoring is important to include starting contributors, at project or at article, where the day-to-day management is conducted (Forte et al., 2009). As Forte et al. (2012, table 1, p. 2) pointed out, relying on McGrath's (1991) typology of (small) group modes and functions, this is exactly what these nested organizations are made for, in addition to production activities support: maintaining group's well being and providing support to members. Finally, the literature on innovative teams (Defélix et al., 2005), not to speak of the literature on virtual teams (Crowston et al., 2010) may present competing explanations of the functioning of these online groups regarding the inclusion of newcomers.

All these put together, it remains that the availability of the Dump files for Wikipedia, of the CVS files for FLOSS projects (from Sourceforge, for instance), makes it possible to do longitudinal studies to evaluate the links developed between people working on the same pieces of knowledge (article/files or sub projects), **thus to measure more precisely the size of the teams and the level of connection between the members.** This requires capacities in data extraction, data mining and especially social network analysis, another field of study where collaboration with colleagues from computer science may allow us to conduct this kind of breakthrough research. The corollary of this study is person's career studies; if I looked quite deeply at individual trajectories and motivations, some questions remain, which can be studied with the same set of data as the team, analyzed at individual level, this time: who the person has worked with, knows, on what kind of piece of knowledge this person has worked, etc.

4.3.3 Participants' competences.

Two points may be of particular interest regarding the involvement and the development of participant's capital.

As already said, what Ostrom and Hess pointed out is that **online communities decreased the boundary between the in and out**. It is easier to find and to join this group than a closed, semi hidden, in-a-word marginalized community of practice such as the witch communities, as showed by Merriam et al. (2003). However, **it doesn't seem to be the case for FLOSS projects**. The explanation may be obvious: as there is no language barrier, it is easier to migrate from one open-source project to another than to migrate from the French Wikipedia to the German one, for instance. An experienced developer can be "hired" to participate to a project, as some editors are "hired" to participate to a thematic project in Wikipedia. Her-raiz et al.'s work (2006) supports this hypothesis, showing that "volunteers tend to follow a step-by-step joining process, while hired developers usually experience a "sudden" integration". **Using data analysis of several projects may make possible to control from this bias and test the importance of the first contribution, identifying people participating in multiple project before surveying them.**

Secondly, **still using the Dumps/ CVS data**, it is possible to identify the trajectory of the people in the community, the specialization in role, and **to see if this specialization is somehow connected to their work**. For instance, it seems to be a two level-leadership in Wikipedia: leaders focused on project management, content-based, where discussion and coordination are closely linked to the level of contribution to the article, with strong effects of socialization, and more global, project level managers, aiming at addressing the unresolved cases. I will try to discover whether this dichotomy is valid, and, to make it understandable even if a bit naïve, if content leaders are also content specialists when project management leaders are project managers.

This closes the loop between organizations and virtual online communities, as this will prove, from another perspective, the connexion between two spaces where people develop competences, and collectively produce knowledge. Is this new organization a new model for the firms, and for the development of the competences of the knowledge producers they hire? This new model of a firm (Baldwin and Von Hippel, 2009) can be discussed from an open innovation literature point of view, as defined by Chesbrough (2003); Chesbrough et al. (2006), something I will do in the conclusion.

Chapter 5

General Conclusion.

My research on online communities in a digitalized knowledge economy goes beyond the simple case of FLOSS; it is a way for me to address the evolution of knowledge production, the discussion on the creative company and on open-innovation as I argued with Julien Pénin (Jullien and Pénin, 2013):

"The FLOSS example, even if it is an extreme case, shows that, to benefit from an open innovation dynamic, the company itself has to be in a process of creating value through service, more or less standard (from insurance to assistance to the use through adaptation to the needs). Beside the "servicial" and the "learning" companies¹, it would create a new type of company, the "creative" company, which "would continuously capitalize on its service relationship to make its cognitive capital and especially its specific service offer evolve" (Dang Nguyen et al., 2010). In those companies, "the entire staff is mobilized not only to better serve customer needs but also to anticipate their needs, to identify relevant information and to make evolve the service offer" (ibid). It would be companies focused on their "dynamic capabilities" defined as "the firm's ability to integrate, build and reconfigure internal and external competences to address rapidly changing environments" (Teece et al., 1997), so, as outlined by Dang Nguyen et al., to "create value faster than competitors." Open innova-

¹"The servicial company creates value on the relationship with the customer. It is rich on the informational level because each client mobilizes a different stock of data and induces a particular know-how [...] This model has a significant potential for value creation, but it is also at risk of spiraling costs linked to a lack of organization management of the operations."

"The learning company assumes that the coordination and organization is a collective work between humans that aggregate their skills and talents for a purpose. The organization must be able to capitalize continuously on their knowledge, share their knowledge and build another set of innovations, that is to say a collective know-how adapted to its environment's changes." (Dang Nguyen et al., 2010, p. 10)

tion would be consubstantial to this type of company, whose competitive advantage is based precisely on its ability to organize the dissemination and exchange of knowledge between product users and to facilitate the use of this product, offering, as a business, guaranteed versions, assistance to the use and adaptation.”

This “creative” company is closed to what Cohendet et al. (2010) called the “innovative” firm, and thus to the organization of its governance (Burger-Helmchen and LLerena, 2008), as, in addition to dealing with creativity, and according to the authors, the governance phase is consistent with O’Mahony and Ferraro’s analysis of a FLOSS community’s governance evolution (O’Mahony and Ferraro, 2007; O’Mahony, 2007). The success of online communities seems also to give credit to Grant’s vision of the firm as an aggregation on individuals bringing their knowledge (1996).

In both case, knowledge workers have to learn how to work virtually in virtual conditions. Those companies are the ones which recruit Télécom Bretagne’s students, and the kind of studies I conduct is of growing importance for the definition of their future jobs, or more explicitly, of their future ways of working. This, if needed, sheds light on the importance for engineering schools to conduct research in social sciences, in addition to traditional engineering sciences research. On the other hand, being at Télécom Bretagne, even if it means having to lose connexion with the evolution of the academic discipline, has some advantages for social science researchers: the contact with participants in such communities is facilitated, and we are more aware of the evolution of technologies or of practices, for instance of the computer/hacker culture. In a sense, we, researchers in social sciences in engineering schools, play the role of the companies investing in open communities to follow the technology. My best example of that being, of course, my involvement in FLOSS studies as early as 1998. The other advantage of this position, as I have illustrated in the future research section, is that collaboration with data science researchers is made easier, and although riskier, it could lead to original results. I am looking forward to keeping on proving this in the future.

Chapter 6

References.

6.1 Global References.

- D. ACEMOGLU** and **J.-S. PISCHKE**, February 1998. Why Do Firms Train? Theory and Evidence. *Quarterly Journal of Economics*, 113(1):79–119.
- P. J. ÅGERFALK** and **B. T. FITZGERALD**, 2007. Outsourcing to an unknown workforce: Exploring opensourcing as a global sourcing strategy. *MIS Quarterly*, 32(2):385–409.
- R. AGUILERA**, **D. RUPP**, **C. WILLIAMS**, and **J. GANAPATHI**, 2007. Putting the S Back in CSR: A Multi-level Theory of Social Change in Organizations. *Academy of Management Review*, 32(3): 836–863.
- R. AGUILERA**, **C. WILLIAMS**, **J. CONLEY**, and **D. RUPP**, 2006. Corporate Governance and Corporate Social Responsibility. A Comparative Analysis of the U.K. and the U.S. *Corporate Governance: An International Review*, 14(3):147–157.
- M. AKRICH**, **M. CALLON**, and **B. LATOUR**, 2006. *Sociologie de la traduction : textes fondateurs*. Les presses des Mines de Paris, Paris.
- Y. AMICHAÏ-HAMBURGER**, **N. LAMDAN**, **R. MADIËL**, and **T. HAYAT**, 2008. Personality characteristics of wikipedia members. *CyberPsychology & Behavior*, 11(6):679–681.
- A. AMIN** and **J. ROBERTS**, 2008. Knowing in action: Beyond communities of practice. *Research Policy*, 37:353–369.

- F. ANDERSSON, M. FREEDMAN, J. HALTIWANGER, J. LANE, and K. SHAW**, 2009. Reaching for the Stars: Who Pays for Talent in Innovative Industries? *Economic Journal*, 119(538):F308–F332.
- D. ANTHONY, S. W. SMITH, and T. WILLIAMSON**, 2009. Reputation and reliability in collective goods: The case of the online encyclopedia Wikipedia. *Rationality and Society*, 21(3):283–306. URL: <http://web.mit.edu/iandeseminar/Papers/Fall2005/anthony.pdf>.
- J. ANTIN and C. CHESHIRE**, 2010. Readers are not free-riders: Reading as a form of participation on Wikipedia. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*, pages 127–130, 2010.
- A. ARORA and A. GAMBARDILLA**, 1994. The changing technology of technical change: general and abstract knowledge and the division of innovative labour. *Research Policy*, 23:523–532.
- K. J. ARROW**, 1962b. Economic welfare and the allocation of resources to invention. In **R. R. NELSON**, editor, *The Rate and Direction of Inventive Activity: Economic and Social Factors*. Princeton University Press, Princeton.
- W. B. ARTHUR**, 1989a. Competing technologies: an overview. In **G. DOSI, C. FREEMAN, R. NELSON, J. SILVERBERG, and L. SOETE**, editors, *Technical Change and Economic Theory*, pages 590–607. Frances Pinter, London.
- W. B. ARTHUR**, 1989b. Competing technologies, increasing returns and lock-in by historical events: The dynamics of allocations under increasing returns to scale. *Economic Journal*, 99:116–131. URL: http://www.santafe.edu/arthur/Papers/Pdf_files/EJ.pdf.
- C. G. AU PLAN**, 2002. Économie du logiciel: renforcer la dynamique française. rapport du groupe présidé par Hugues Rougier, Rapporteur Général: Jean-Michel Dalle, Rapporteur: Sylvie Bénard.
- N. AURAY, C. POUDAT, and P. PONS**, 2007. Democratizing Scientific Vulgarization. The Balance between Cooperation and Conflict in French Wikipedia. *Observatorio (OBS*) Journal*, 3:185–199.
- L. BACH, P. COHENDET, J. PÉNIN, and L. SIMON**, 2010. Creative industries and the ipr dilemma between appropriation and creation: some insights from the videogame and music industries. *Management international / International Management / Gestión Internacional*, 14(3):59–72.

- C. Y. BALDWIN** and **E. A. VON HIPPEL**, November 2009. Modeling a Paradigm Shift: From Producer Innovation to User and Open Collaborative Innovation. *Harvard Business School Finance Working Paper No. 10-038*; *MIT Sloan Research Paper No. 4764-09*. URL: <http://ssrn.com/abstract=1502864>.
- K. BALKA**, **C. RAASCH**, and **C. HERSTATT**, 2009. Open source enters the world of atoms: A statistical analysis of open design. *First Monday*, 14(11).
- F. BARCELLINI**, **F. DÉTIENNE**, and **J.-M. BURKHARDT**, July 2008. User and developer mediation in an open source software community: Boundary spanning through cross participation in online discussions. *International Journal of Human-Computer Studies*, 66:558–570.
- S. BEAUD**, 1996. L’usage de l’entretien en sciences sociales. plaidoyer pour l’"entretien ethnographique". *Politix*, 9(35):226–257.
- S. BEAUD** and **F. WEBER**, 2003. *Guide de l’enquête de terrain: produire et analyser des données ethnographiques*. La Découverte, Paris.
- H. S. BECKER**, 1960. Notes on the concept of commitment. *American Journal of Sociology*, 66(1): 32–40.
- H. S. BECKER**, 1963. *Outsiders. Studies in the Sociology of Deviance*. Free Press.
- Y. BENKER** and **H. NISSENBAUM**, 2006. Commons-based peer production and virtue. *The Journal of Political Philosophy*, 14(4):394–419.
- D. BERTAUX**, 2010. *Le récit de vie. L’enquête et ses méthodes*. col. 128, Armand-Colin, 3 edition.
- J. BESSEN**, 2005. Open Source Software: Free Provision of Complex Public Goods. Technical report, Research on Innovation. URL: <http://www.researchoninnovation.org/opensrc.pdf>.
- M. D. BESTEN** and **J.-M. DALLE**, 2008. Keep it simple: A companion for simple wikipedia? *Industry & Innovation*, 15(2):169–178.
- S. BHATTACHARYA** and **S. GURIEV**, 2006. Patents vs. trade secrets: Knowledge licensing and spillover. *Journal of the Economic Association*, 4(6):1112–1147.
- A. BLANCHET** and **A. GOTMAN**, 2010. *L’enquête et ses méthodes*. col. 128, Armand-Colin, 2 edition.

- A. BONACCORSI** and **C. ROSSI**, July 2003. Why open source software can succeed. *Research Policy*, 32(7):1243–1258.
- S. BORZILLO**, **S. AZNAR**, and **A. SCHMITT**, February 2011. A journey through communities of practice: How and why members move from the periphery to the core. *European Management Journal*, 29(1):25–42.
- A. BRANDENBURGER** and **B. NALEBUFF**, 1996. *Co-opetition: A Revolution Mindset That Combines Competition and Cooperation... The Game Theory Strategy That's Changing the Game of Business*. Doubleday Book, New York.
- U. BRANDES** and **J. LERNER**, March 2008. Visual analysis of controversy in user-generated encyclopedias. *Information Visualization*, 7:34–48.
- E. BROUSSEAU** and **N. CURIEN**, 2007. *Internet and Digital Economics; Principles, Methods and Applications*. Cambridge University Press.
- S. L. BRYANT**, **A. FORTE**, and **A. BRUCKMAN**, 2005. Becoming Wikipedian: Transformation of Participation in a Collaborative Online Encyclopedia. In *proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*. ACM, 2005.
- E. BRYNJOLFSSON**, 2004. Productivity's technology iceberg. the real story behind america's productivity revival is the way companies have combined digital technologies with digital organization. *Technology Review*, March 10. URL: <http://www.technologyreview.com/web/13511/>.
- E. BRYNJOLFSSON** and **L. M. HITT**, 2003. Computing productivity: Firm-level evidence. *Review of Economics and Statistics*, 85(4):793–808.
- G. BUONANNO**, **P. FAVERIO**, **F. PIGNI**, **A. RAVARINI**, **D. SCIUTO**, and **M. TAGLIAVINI**, 2005. Factors affecting ERP system adoption: A comparative analysis between SMEs and large companies. *Journal of Enterprise Information Management*, 18(4):384–426.
- A. BURETH**, **M. MÜLLER**, **J. PÉNIN**, and **S. WOLF**, 4th quarter 2007. Brevet, innovation modulaire et collaboration. *Revue d'économie industrielle*, 120. URL: <http://rei.revues.org/index3123.html>.

- T. BURGER-HELMCHEN** and **P. LLERENA**, 2008. A case study of a creative start-up: governance, communities and knowledge management. *Journal of Innovation Economics*, 2:125–146.
- L. BURIOL**, **C. CASTILLO**, **D. DONATO**, **S. LEONARDI**, and **S. MILLOZZI**, December 2006. Temporal Evolution of the Wikigraph. In *Proceedings of Web Intelligence*, pages 45–51, Hong Kong. IEEE CS Press. URL: http://www.chato.cl/papers/buriol_2006_temporal_analysis_wikigraph.pdf.
- K. L. CALLIS**, **L. R. CHRIST**, **J. RESASCO**, **D. W. ARMITAGE**, **J. D. ASH**, **T. T. CAUGHLIN**, **S. F. CLEMMENSEN**, **S. M. COPELAND**, **T. J. FULLMAN**, **R. L. LYNCH**, **C. OLSON**, **R. A. PRUNER**, **E. H. VIEIRA-NETO**, **R. WEST-SINGH**, and **E. M. BRUNA**, 2009. Improving Wikipedia: educational opportunity and professional responsibility. *Trends in Ecology & Evolution*, 24(4):177 – 179.
- M. CALLON**, 1994. Is science a public good? fifth mullins lecture. *Science, Tehcnology and Human Value*, 14-4.
- M. CALLON**, 1999. Le réseau comme forme émergente et comme modalité de coordination: le cas des interactions stratégiques entre firmes industrielles et laboratoires académiques. In **CALLON ET AL. (1999)**, editor, *Réseau et coordination*, pages 13–64.
- M. CALLON**, **P. COHENDET**, **N. CURIEN**, **J.-M. DALLE**, **F. EYMARD-DUVERNAY**, **D. FORAY**, and **E. SCHENK**, editors, 1999. *Réseau et coordination*. Economica.
- M. CALLON**, **J. LAW**, and **A. RIP**, 1986. *Mapping the Dynamics of Science and Technology*. MacMillan.
- K. CARILLO** and **C. OKOLI**, 2011. Generating quality open content: A functional group perspective based on the time, interaction, and performance theory. *Information & Management*, 48(6):208–219.
- R. A. CARVALHO**, 2009. Free/open source enterprise resources planning. In **J. GUPTA**, editor, *Handbook of Enterprise Systems*, pages 99–111. Information Science Reference.
- S.-I. CHANG**, **S.-Y. HUNG**, **D. YEN**, and **P.-J. LEE**, 2010. Critical factors of erp adoption for small- and medium-sized enterprises: An empirical study. *Journal of Global Information Management*, 18 (3):82–106.

- A. CHARNES, W. COOPER, and R. E.**, 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2:429–444.
- H. CHESBROUGH**, 2003. *Open Innovation: The New Imperative for Creating and Profiting from Technology*. Harvard Business School Press, Boston, MA.
- H. CHESBROUGH**, 2006. Open innovation: a new paradigm for understanding industrial innovation. In **CHESBROUGH ET AL. (2006)**, editor, *Open Innovation: Researching a New Paradigm*, pages 1–12.
- H. CHESBROUGH, W. VANHAVERBEKE, and J. WEST**, editors, 2006. *Open Innovation: Researching a New Paradigm*. Oxford University Press, Oxford.
- H. CHO, M. CHEN, and S. CHUNG**, 2010a. Testing an integrative theoretical model of knowledge-sharing behavior in the context of wikipedia. *Journal of the American Society for Information Science and Technology*, 61(6):1198–1212.
- M. CLÉMENT-FONTAINE**, 2002. Les licences libres. analyse juridique. In **(JULLIEN ET AL., 2002B)**, editor, *Nouveaux modèles économiques, nouvelle économie du logiciel*, pages 163–177.
- M. CLÉMENT-FONTAINE**, 2009. L’œuvre libre. *Juris-classeur PLA*, fasc 1975.
- M. D. COHEN, J. G. MARCH, and J. P. OLSEN**, 1972. A garbage can model of organizational choice. *Administrative Science Quarterly*, 17(1):1–25.
- W. M. COHEN and D. A. LEVINTHAL**, 1989. Innovation and learning: The two faces of r&d. *Economic Journal*, 99:569–596.
- W. M. COHEN and D. A. LEVINTHAL**, 1990. Absorptive capacity, a new perspective of learning and innovation. *Administrative Science Quarterly*, 35:128–152.
- P. COHENDET, F. CRÉPLET, and O. DUPOUET**, 2001. Interactions between epistemic communities and communities of practice as a mechanism of creation and diffusion of knowledge. In **J.-B. ZIMMERMANN and A. KIRMAN**, editors, *Interaction and Market Structure*. Springer, Londres.
- P. COHENDET, F. CREPLET, and O. DUPOUET**, 2002. Innovation organisationnelle, communautés de pratique et communautés épistémiques: le cas de linux. *Revue française de gestion*, to be published.

- P. COHENDET, P. LLERENA, and L. SIMON**, 2010. The innovative firm: nexus of communities and creativity. *Revue d'économie industrielle*, 129-130:139–170.
- S. COMINO, F. M. MANENTI, and A. ROSSI**, 2011. Public intervention for free/open source software. *Revue d'économie industrielle*, 135:89–108.
- W. W. COOPER, L. M. SEIFORD, and J. ZHU**, editors, 2011. *Handbook on Data Envelopment Analysis*. Springer, 2nd edition.
- M. CORIS**, 2006. Les consortiums ouverts: une nouvelle approche pour la standardisation des logiciels? *Revue d'économie industrielle*, 116(4e trim):105–126.
- J. CRÉMER, C. D'ASPREMONT, and L.-A. GÉRARD-VARET**, August 1990. Incentives and the existence of pareto-optimal revelation mechanisms. *Journal of Economic Theory*, 51(2):233–254.
- K. CROWSTON**, June 2011. Lessons from volunteering and free/libre open source software development for the future of work. In *Proceedings of Researching The Future in Information Systems: IFIP Working Group 8.2 Working Conference, Future IS*, page 215, Turku, Finland. Springer. URL: http://crowston.syr.edu/system/files/ifipwg82paper_final.pdf.
- K. CROWSTON, R. HECKMAN, and N. MISIOLEK**, December 2010. Leadership in self-managing virtual teams. *Syracuse University School of Information Studies working paper*.
- K. CROWSTON and J. HOWISON**, February 2005. The social structure of Free and Open Source Software development. *First Monday*, 10(2). URL: <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1207/1127>.
- K. CROWSTON, J. HOWISON, and H. ANNABI**, 2006. Success in free and open source software development: Theory and measures. *Software Process Improvement and Practice*, 11:123–148.
- M. CROZIER and E. FRIEDBERG**, 1997. *L'acteur et le système*. Seuil, Paris.
- M. CUSUMANO**, 2004. *The Business of software: what every manager, programmer, and entrepreneur must know to thrive and survive in good times and bad*. Free Press, New York.
- L. DAHLANDER and M. G. MAGNUSSON**, 2005. Relationships Between Open Source Software Companies and Communities: Observations from Nordic Firms. *Research Policy*, 34:481–493.

- L. DAHLANDER** and **M. G. MAGNUSSON**, 2008. How do firms make use of open source communities? *Long Range Planning*, 41 (2008):629–649.
- F. DAMANPOUR**, 1991. Organizational innovation: A meta-analysis of effects of determinants and moderators. *The Academy of Management Journal*, 34(3):555–590.
- G. DANG NGUYEN** and **C. GENTHON**, 2006. Les perspectives du secteur des Tic en Europe (Perspectives for the European ICT Industry). *Marsouin Working Paper*, 4-2006.
- G. DANG NGUYEN**, **Y. LERAY**, and **O. MÉVEL**, 2010. Enseignements et prospective des pratiques des entreprises de services. In *Proceedings. 10e rencontre sur la prospective des métiers: quel management demain?*, 2010.
- É. DARMON**, **T. LE TEXIER**, and **D. TORRE**, 2011. Commercial or Open Source Software? Winner-Takes-All Competition, Partial Adoption and Efficiency. *Revue d'économie industrielle*, 135:109–140.
- É. DARMON** and **D. TORRE**, 2009. Open source and commercial software platforms: Is coexistence a temporary or a sustainable outcome? *International Journal of Open Source Software & Processes*, 1 (2):67–80.
- F. D. DAVIS**, 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3):319–340.
- J. DE BANDT**, 1995. *Services aux entreprises: informations, produits, richesses*. Economica, Paris.
- C. DEDE**, May 2008. A seismic shift in epistemology. *EDUCAUSE Review*, 43:80–81. URL: <http://www.educause.edu/EDUCAUSE+Review/EDUCAUSEReviewMagazineVolume43/ASeismicShiftinEpistemology/162892>.
- J. DEDRICK**, **V. GURBAXANI**, and **K. L. KRAEMER**, March 2003. Information technology and economic performance: A critical review of the empirical evidence. *ACM Computing Surveys*, 35(1): 1–28.
- J. DEDRICK** and **J. WEST**, jan. 2004. An exploratory study into open source platform adoption. In *Proceedings of the 37th Annual Hawaii International Conference on System Sciences*, page 10 pp., jan. 2004.

- C. DEFÉLIX, V. CHANAL, and B. GALEY**, 2005. Les personnes innovantes dans les entreprises doivent-elles faire l'objet d'une GRH spécifique ? Une étude exploratoire. *Gestion 2000*, mars-avril(2):99–113.
- W. H. DELONE and E. R. MCLEAN**, 1992. Information systems success: The quest for the dependent variable. *Information Systems Research*, 3(1):60–95.
- W. H. DELONE and E. R. MCLEAN**, 2002. Information systems success revisited. In *Proceedings of the 35th Hawaii International Conference on System Sciences*, Waikola, Hawaii. HICSS-35. URL: <http://csdl2.computer.org/comp/proceedings/hicss/2002/1435/08/14350238.pdf>.
- W. H. DELONE and E. R. MCLEAN**, 2003. The delone and mclean model of information systems success: a ten-year update. *Journal of Management Information Systems*, 19(4):9–30.
- F. DELTOUR and J. TRÉMENBERT**, 2010. L'adoption de progiciels de gestion intégrée dans les pme: des débuts prometteurs. Technical report. URL: <http://www.marsouin.org/spip.php?article379>.
- P. DENNING, J. HORNING, D. PARNAS, and L. WEINSTEIN**, December 2005. Wikipedia risks. *Commun. ACM*, 48:152–152.
- N. K. DENZIN and Y. S. LINCOLN**, 2008. *Strategies of qualitative inquiry*. Sage, Thousand Oaks, CA, 3 edition.
- N. K. DENZIN and Y. S. LINCOLN**, editors, 2011. *The Sage Handbook of Qualitative Research*. Sage, Thousand Oaks, CA, 4 edition.
- K. EDWARDS**, 2001. Epistemic Communities, Situated Learning and Open Source Software Development. In *Cultures and the Practice of Interdisciplinarity' Workshop at NTNU*, page 24, 2001.
- H. EIJKMAN**, 2008. Web 2.0 as a non-foundational network-centric learning space. *Campus-Wide Information Systems*, 25(2):93 – 104.
- D. FALLIS**, 2008. Toward an epistemology of wikipedia. *Journal of the American Society for Information Science and Technology*, 59(10):1662–1674.
- Y. FANG and D. NEUFELD**, April 2009. Understanding Sustained Participation in Open Source Software Projects. *Journal on Management Information Systems*, 25(4):9–50.

- M. J. FARELL**, 1957. The measurement of productive efficiency. *Journal of the Royal Statistical Society, Series A*, 120(3):250–290.
- J. FARRELL**, 1989. Standardization and intellectual property. *Jurimetrics Journal*, 30:35.
- J. FARRELL** and **G. SALONER**, 1985. Standardisation, compatibility and innovation. *Rand Journal of Economics*, 16:70–83.
- J. FELLER**, **R. FITZGERALD**, **S. HISSAM**, and **R. K. LAKHANI**, editors, 2005. *Perspectives on free and open source software*. MIT Press, New York.
- M. M. FERRITER**, March 2009. "Arguably the Greatest": Sport Fans and Communities at Work on Wikipedia. *Sociology of Sport Journal*, 26(1):127–154.
- R. G. FICHMAN**, 1992. Information technology diffusion: A review of empirical research. In *Proceedings of the Thirteenth International Conference on Information Systems*, pages 195–206, 1992.
- R. G. FICHMAN** and **C. F. KEMERER**, 1993. Adoption of software engineering process innovations: The case of object orientation. *Sloan Management Review*, 34(2):7–22.
- T. FIEDLER**, 2008. The web's pathway to accuracy. *Nieman Reports*, 62.
- B. FITZGERALD**, 2006. The transformation of open source software. *MIS Quarterly*, 30(3):587–598.
- C. FLAHERTY MANCHESTER**, October 2008. The Effect of Tuition Reimbursement on Turnover: A Case Study Analysis. In **S. BENDER**, **J. LANE**, **K. SHAW**, **F. ANDERSSON**, and **T. VON WACHTER**, editors, *The Analysis of Firms and Employees: Quantitative and Qualitative Approaches*, pages 197–228. University of Chicago Press.
- B. FLYVBJERG**, 2011. Case study. In (**DENZIN AND LINCOLN, 2011**), editor, *The Sage Handbook of Qualitative Research*, chapter 17, pages 301–316.
- P. K.-F. FONG** and **R. P. BIUK-AGHAI**, 2010. What did they do? Deriving high-level edit histories in Wikis. In *Proceedings of the 6th International Symposium on Wikis and Open Collaboration*, WikiSym '10, pages 2:1–2:10, New York, NY, USA. ACM.
- D. FORAY** and **M. CASSIER**, 2001a. Économie de la connaissance: le rôle des consortiums de haute technologie dans la production de bien public. *Économie et Prévision*, 4/5(150-151):107–122.

- D. FORAY** and **M. CASSIER**, 2001b. Public Knowledge, Private Property and the Economics of High-tech Consortia. *Economics of Innovation and New Technology*, 11(2):123–132.
- D. FORAY**, **S. THORON**, and **J.-B. ZIMMERMANN**, 2007. Open software: Knowledge openness and cooperation in cyberspace. In (**BROUSSEAU AND CURIEN**, 2007), editor, *Internet and Digital Economics; Principles, Methods and Applications*, pages 368–390.
- D. FORAY** and **J.-B. ZIMMERMANN**, October 2001. L'économie du logiciel libre: organisation coopérative et incitation à l'innovation. *Revue économique*, 52:77–93. Special Issue, hors série: économie d'Internet, sous la direction d'É. Brousseau et N. Curien.
- A. FORTE** and **A. BRUCKMAN**, 2005. Why do people write for Wikipedia? Incentives to contribute to open-content publishing. *working paper*.
- A. FORTE**, **N. KITUR**, **V. LARCO**, **H. ZHU**, **A. BRUCKMAN**, and **R. E. KRAUT**, 2012. Coordination and beyond: social functions of groups in open content production. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work, CSCW '12*, pages 417–426, New York, NY, USA. ACM.
- A. FORTE**, **V. LARCO**, and **A. BRUCKMAN**, 2009. Decentralization in wikipedia governance. *Journal of Management Information Systems*, 26(1):49–72.
- D. FRÉARD**, **A. DENIS**, **F. DÉTIENNE**, **M. BAKER**, **M. QUIGNARD**, and **F. BARCELLINI**, 2010. The role of argumentation in online epistemic communities: the anatomy of a conflict in wikipedia. In *Proceedings of the 28th Annual European Conference on Cognitive Ergonomics, ECCE '10*, pages 91–98, New York, NY, USA. ACM.
- P. GEROSKI**, 1995. What do we know about entry? *International Journal of Industrial Organization*, 13(4):421 – 440.
- J. GILES**, December 2005. Internet encyclopaedias go head to head. *Nature*, 438:900–901.
- E. GLEAVE**, **H. T. WELSER**, **T. M. LENTO**, and **M. A. SMITH**, January 2009. A conceptual and operational definition of 'social role' in online community. In *Proceedings, Hawaii. 42nd Hawaii International Conference on System Sciences*, IEEE.

- R. GLOTT, P. SCHMIDT, and R. GHOSH**, 2010b. Analysis of wikipedia survey data. topic: Quality of wikipedia content. Technical report, UNU-MERIT. URL: http://www.wikipediasurvey.org/docs/Wikipedia_Overview_15March2010-FINAL.pdf.
- R. GLOTT, P. SCHMIDT, and R. GHOSH**, 2010c. Wikipedia survey – overview of results. Technical report, UNU-MERIT. URL: http://www.wikipediasurvey.org/docs/Wikipedia_Overview_15March2010-FINAL.pdf.
- G. GORMAN**, 2007. A tale of information ethics and encyclopaedias; or, is it wikipedia/it just another internet scam? *Online Information Review*, 31(3):273–276. URL: <http://www.ingentaconnect.com/content/mcb/264/2007/00000031/00000003/art00001>.
- R. GRANT**, 1996. Toward a knowledge-based theory of the firm. *Strategic Management Journal*, 17 (Winter):109–122.
- J. R. HACKMAN**, 1987. The design of work teams. In **J. W. LORSCH**, editor, *Handbook of Organizational Behavior*, pages 315–342. Prentice Hall, Englewood Cliffs, NJ.
- M. HADDARA and O. ZACH**, January 2011. Erp systems in smes: A literature review. In *44th Hawaii International Conference on System Sciences (HICSS)*, pages 1–10, January 2011.
- C. A. HAIGH**, 2011. Wikipedia as an evidence source for nursing and healthcare students. *Nurse Education Today*, 31(2):135 – 139.
- A. HALFAKER, A. KITTUR, and J. RIEDL**, October 2011. Don’t bite the newbies: How reverts affect the quantity and quality of wikipedia work. In *Proceedings of the 7th International Symposium on Wikis and Open Collaboration, WikiSym ’11*, New York, NY, USA. ACM.
- H. HAMILTON, J. A. NICKERSON, H. OWAN, R. POLLACK, R. KOENKER, B. FITZENBERGER, A. RAO, G. MACDONALD, and S. LEVITT**, 2003. Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation. *Journal of Political Economy*, pages 465–497.
- N. HARA, P. SHACHAF, and K. F. HEW**, 2010. Cross-cultural analysis of the Wikipedia community. *Journal of the American Society of Information Science and Technology*, 61(10):2097–2108.

- J. HECKMAN**, 1979. Sample selection bias as a specification error. *Econometrica*, 47(1):153–161.
- A. HEMETSBERGER** and **C. REINHARDT**, 2009. Collective Development in Open-Source Communities: An Activity Theoretical Perspective on Successful Online Collaboration. *Organization Studies*, 30:987–1008.
- J. HENKEL**, 2006a. The jukebox mode of innovation - a model of commercial open source development. *DRUID Working Paper*, 06-25.
- J. HENKEL**, 2006b. Selective revealing in open innovation processes: The case of embedded linux. *Research Policy*, 35(7):953 – 969.
- M. HERNANDEZ**, April 2012. Toward an Understanding of the Psychology of Stewardship. *The Academy of Management Review (AMR)*, 37(2):172–193.
- I. HERRAIZ**, **G. ROBLES**, **J. J. AMOR**, **T. ROMERA**, and **J. M. GONZÁLEZ-BARAHONA**, 2006. The processes of joining in global distributed software projects. In *GSD '06: Proceedings of the 2006 international workshop on Global software development for the practitioner*, pages 27–33, New York, NY, USA. ACM.
- G. HERTEL**, **S. NIEDNER**, and **S. HERRMANN**, 2003. Motivation of software developers in open source projects: an internet-based survey of contributors to the linux kernel. *Research Policy*, 32:1159–1177.
- C. HESS** and **E. OSTROM**, 2006a. A Framework for Analyzing the Knowledge Commons. In (**HESS AND OSTROM, 2006C**), editor, *Understanding Knowledge as a Commons. From Theory to Practice*, pages 41–81.
- C. HESS** and **E. OSTROM**, 2006b. Introduction: An Overview of the Knowledge Commons. In (**HESS AND OSTROM, 2006C**), editor, *Understanding Knowledge as a Commons. From Theory to Practice*, pages 3–26.
- C. HESS** and **E. OSTROM**, editors, december 2006c. *Understanding Knowledge as a Commons. From Theory to Practice*. MIT Press.
- S. HO**, **P. TING**, **D. BAU**, and **C. WEI**, September 2011. Knowledge-sharing intention in a virtual community: A study of participants in the chinese wikipedia. *Cyberpsychology Behavior and social networking*, 14(9):541–545.

- G. HOFSTEDE**, 1991. *Cultures and organizations: Software of the mind*. McGraw-Hill, London.
- J. HOPE**, 2008. *Biobazaar: The Open Source Revolution and Biotechnology*. Haward University Press.
- F. HORN**, 1999b. L'importance du logiciel libre dans l'amélioration de l'efficience des logiciels. *Terminal*, 80/81:119–148. Special Issue, *Le logiciel libre*.
- F. HORN**, 2004. *L'économie des logiciels*. repères, la Découverte.
- M. IANSITI** and **G. L. RICHARDS**, 2006. The business of free software: Enterprise incentives, investment, and motivation in the open source community. *Harvard Business School Working Paper Serie*, 07-028.
- T. IBA**, **K. NEMOTO**, **B. PETERS**, and **P. A. GLOOR**, 2010. Analyzing the creative editing behavior of wikipedia editors through dynamic social network analysis. In *Procedia - Social and Behavioral Sciences*, volume 2, pages 6441–6456, 2010.
- C. JENSEN** and **W. SCACCHI**, 2007. Role Migration and Advancement Processes in OSSD Projects: A Comparative Case Study. In *ICSE 07: Proceedings of the 29th international conference on Software Engineering*, pages 364–374, Washington, DC, USA. IEEE Computer Society.
- B. JOHANSSON** and **F. SUDZINA**, 2008. ERP systems and open source: An initial review and some implications for SMEs. *Journal of Enterprise Information Management*, 21(6):649–658.
- P. JOLY** and **B. HERVIEU**, December 2003. La marchandisation du vivant. *Futuribles*, 292:5–30.
- M. KAMATA**, **D. KATO**, **K. KUNIEDA**, and **K. YAMADA**, 2010. Web community contributor's motivation: Japanese wikipedia case study. In *Proc. of the IADIS Int. Conf. Collaborative Technologies 2010, Proc. of the IADIS Int. Conf. Web Based Communities 2010, Part of the MCCSIS 2010*, pages 29–33, 2010.
- M. L. KATZ** and **C. SHAPIRO**, 1986b. Technology adoption in the presence of network externalities. *Journal of Political Economy*, 75/4.
- M. L. KATZ** and **C. SHAPIRO**, spring 1994. Systems competition and network effects. *Journal of Economic Perspectives*, 8(4):93–115.

- B. A. KITCHENHAM**, 2002. The question of scale economies in software - why cannot researchers agree? *Information and Software Technology*, 44(1):13 – 24.
- A. KITTUR, E. H. CHI, B. A. PENDLETON, B. SUH, and T. MYTKOWICZ**, 2007a. Power of the few vs. wisdom of the crowd: Wikipedia and the rise of the bourgeoisie. In *CHI 2007*, San Jose, CA, USA.
- A. KITTUR and R. E. KRAUT**, 2008. Harnessing the wisdom of crowds in wikipedia: quality through coordination. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work, CSCW '08*, pages 37–46, New York, NY, USA. ACM.
- A. KITTUR, B. PENDLETON, and R. E. KRAUT**, 2009. Herding the cats: The influence of groups in coordinating peer production. In *Proceedings of the 5th International Symposium on Wikis and Open Collaboration, WiKiSym 2009*, 2009.
- A. KITTUR, B. SUH, B. PENDLETON, and E. H. CHI**, 2007b. He says, she says: Conflict and coordination in wikipedia. In *CHI 2007*, San Jose, CA, USA.
- S. KOCH**, 2008a. Effort modeling and programmer participation in open source software projects. *Information Economics and Policy*, 20(4):345–355.
- S. KOCH**, 2008b. Measuring the efficiency of free and open source software projects using data envelopment analysis. In **S. SOWE, I. STAMELOS, and I. SAMOLADAS**, editors, *Emerging Free and Open Source Software Practices*, pages 25–44. Hershey, PA: IGI.
- S. KOCH**, 2009. Exploring the effects of sourceforge.net coordination and communication tools on the efficiency of open source projects using data envelopment analysis. *Empirical Software Engineering*, 14:397–417.
- S. KOCH**, 2011. Organisation of work in open source projects: expended effort and efficiency. *Revue d'économie industrielle*, 136:17–38.
- B. M. KOGUT and A. METIU**, August 2001. Open source software development and distributed innovation. *Oxford Review of Economic Policy*, 17(2):248–264.
- R. KOPPL and R. LANGLOIS**, 2001. Organization and language games. *Journal of Management and Governance*, 5(3-4):287–305.

- J.-L. LACOLLEY, T. LOILIER, and A. TELLIER**, June 2007. La prise de décision dans les équipes de la communauté des logiciels libres: Faut-il mettre le bazar dans la poubelle? In *XVIe Conférence Internationale de Management Stratégique*, Montréal.
- K. LAKHANI and E. VON HIPPEL**, 2003. How open source software works: Free user to user assistance. *Research Policy*, 32:923–943. URL: <http://opensource.mit.edu/papers/lakhanivonhippelusersupport.pdf>.
- K. LAKHANI and R. WOLF**, 2005. Why hackers do what they do: Understanding motivation and effort in free/open source software projects. In *Perspectives on free and open source software*, pages 3–22. in Feller et al. (2005) edition.
- S. K. LAM, A. UDUWAGE, Z. DONG, S. SEN, D. R. MUSICANT, L. TERVEEN, and J. RIEDL**, October 2011. WP:Clubhouse? An Exploration of Wikipedia’s Gender Imbalance. In *Proceedings of the 7th International Symposium on Wikis and Open Collaboration*, WikiSym ’11, New York, NY, USA. ACM.
- E. LAMY and T. SHINN**, 2006a. L’autonomie scientifique face à la mercantilisation. formes d’engagement entrepreneurial des chercheurs en france. *Actes de la recherche en sciences sociales*, 4(164):23–50.
- R. N. LANGLOIS and P. L. ROBERTSON**, August 1992. Networks and innovation in a modular system: Lessons from the microcomputer and stereo component industries. *Research Policy*, 21(4):297–313.
- D. LANIADO and R. TASSO**, 2011. Co-authorship 2.0: Patterns of collaboration in Wikipedia. In *Proceedings of the 22nd ACM conference on Hypertext and hypermedia*, HT ’11, pages 201–210, New York, NY, USA. ACM.
- B. LATOUR and S. WOOLGAR**, 1979. *Laboratory Life: The Social Construction of Scientific Facts*. Sage Publications, Beverly Hills.
- J. LAVE and E. WENGER**, 1991. *Situated learning: legitimate peripheral participation*. Cambridge University Press, New York.
- E. LAZEAR and P. OYER**, 2013. Personnel Economics. In **R. GIBBONS and D. J. ROBERTS**, editors, *Handbook of Organizational Economics*, page to be published. Princeton University Press.

- S.-Y. T. LEE, H.-W. KIM, and S. GUPTA**, 2009. Measuring open source software success. *Omega*, 37 (2):426 – 438.
- J. LERNER and J. TIROLE**, June 2002. Some simple economics of open source. *Journal of Industrial Economics*, 50:197–234.
- D. LEWANDOWSKI and U. SPREE**, 2011. Ranking of Wikipedia articles in search engines revisited: Fair ranking for reasonable quality? *Journal of the American Society for Information Science and Technology*, 62(1):117–132.
- Q. LI, R. HECKMAN, K. CROWSTON, J. HOWISON, E. E. ALLEN, and U. Y. ESERYEL**, 2008. Decision making paths in self-organizing technology-mediated distributed teams. In *Proceedings of the International Conference on Open Source Systems*, Paris, France, 14-17 December.
- C. LIANG, C. CHEN, and Y. HSU**, 2008. The participation motivation and work styles of the administrators for chinese wikipedia. *Journal of Educational Media and Library Science*, 46(1):81–110.
- H. LIANG CHEN**, 2010. The perspectives of higher education faculty on wikipedia. *The Electronic Library*, 28(3):361–373.
- A. LIH**, 2004. Wikipedia as participatory journalism: Reliable sources? metrics for evaluating collaborative media as a news resource. In *5th International Symposium on Online Journalism*, 2004.
- P. MAGNUS**, 2006. Epistemology and the wikipedia. *North American Computing and Philosophy Conference Troy*.
- D. MAHR and A. LIEVENS**, 2012. Virtual lead user communities: Drivers of knowledge creation for innovation. *Research Policy*, 41(1):167 – 177.
- I. MAIGNAN and D. RALSTON**, 2002. Corporate Social Responsibility in Europe and the U.S.: Insights from Businesses Self-presentations. *Journal of International Business Studies*, 33(3):497–514.
- T. MAILLART, D. SORNETTE, S. SPAETH, and G. VON KROGH**, Nov 2008. Empirical Tests of Zipf’s Law Mechanism in Open Source Linux Distribution. *Physical Review Letters*, 101:218701.
- J. G. MARCH**, 1981. Footnotes to change. *Administrative Science Quarterly*, 26(4):563–577.

- A. M. MARINO** and **J. ZÁBOJNÁK**, 2008. Work-Related Perks, Agency Problems, and Optimal Incentive Contracts. *The RAND Journal of Economics*, 39(2):565–585. URL: <http://www.jstor.org/stable/25474383>.
- C. MARKIDES** and **P. GEROSKI**, 2005. *Fast Second: How Smart Companies Bypass Radical Innovation To Enter And Dominate New Markets*. Jossey-Bass.
- G. MARWELL** and **P. OLIVER**, 1993. *The Critical Mass in Collective Action: A Micro-Social Theory*. Cambridge University Press, Cambridge.
- A. MAS** and **E. MORETTI**, 2009. Peers at Work. *American Economic Review*, 99(1):112–145.
- J. MATEOS-GARCIA** and **W. E. STEINMUELLER**, June 2003. Applying the Open Source Development Model to Knowledge Work. SPRU Electronic Working Paper Series 94, University of Sussex, SPRU - Science and Technology Policy Research. URL: <http://ideas.repec.org/p/sru/ssewps/94.html>.
- J. MATEOS-GARCIA** and **W. E. STEINMUELLER**, 2008. The institutions of open source software: Examining the Debian community. *Information Economics and Policy*, 20(4):333–344.
- J. MCGRATH**, 1991. Time, Interaction, and Performance (TIP): A Theory of Groups. *Small Group Research*, 22(2):147–174.
- S. B. MERRIAM**, **B. COURTENAY**, and **L. BAUMGARTNER**, 2003. On Becoming a Witch: Learning in a Marginalized Community of Practice. *Adult Education Quarterly*, 53(3):170–188.
- MERRIAM-WEBSTER**, 2009. *Online Dictionary*. URL: <http://www.merriam-webster.com/dictionary/>.
- P. MEYER** and **G. PONTIÈRE**, 2011. Eliciting preferences on multiattribute societies with a Choquet integral. *Computational economics*, 37(2):133–168.
- L. MORGAN** and **P. FINNEGAN**, 2007. How Perceptions of Open Source Software Influence Adoption: An Exploratory Study. In *Proceedings of the 15th European Conference on Information Systems*, 2007. URL: <http://is2.lse.ac.uk/asp/aspecis/20070050.pdf>.

- D. R. MUSICANT, Y. REN, J. A. JOHNSON, and J. RIEDL**, October 2011. Mentoring in Wikipedia: A Clash of Cultures. In *Proceedings of the 7th International Symposium on Wikis and Open Collaboration*, WikiSym '11, New York, NY, USA. ACM.
- O. NOV**, November 2007. What motivates wikipedians? *Communications of the ACM*, 50:60–64.
- V. O'DAY, A. ADLER, A. KUCHINSKY, and A. BOUCH**, 16-20 September 2001. When worlds collide: Molecular biology as interdisciplinary collaboration. In **W. PNNZ, M. JARKE, Y. ROGERS, K. SCHMIDT, and V. WULF**, editors, *Proceedings*, pages 399–418, Bonn, Germany. the Seventh European Conference on Computer-Supported Cooperative Work, Kluwer Academic.
- OECD**, 1996. The knowledge-based Economy. Technical Report (96)102.
- C. OKOLI**, 2009. A Brief Review of Studies of Wikipedia in Peer-Reviewed Journals. *International Conference on the Digital Society*, 0:155–160.
- C. OKOLI, M. MEHDI, M. MESGARI, F. Å. NIELSEN, and A. LANAMÄKI**, 2012. The People's Encyclopedia Under the Gaze of the Sages: A Systematic Review of Scholarly Research on Wikipedia. Technical report. URL: <http://ssrn.com/paper=2021326>.
- C. OKOLI and W. OH**, April 2007. Investigating recognition-based performance in an open content community: A social capital perspective. *Information & Management*, 44:240–252.
- P. OLIVER, G. MARWELL, and R. TEIXEIRA**, November 1985. A theory of critical mass interdependence, group heterogeneity, and the production of collective action. *American Journal of Sociology*, 91(3):522–556.
- F. X. OLLEROS**, 2008. Learning to trust the crowd: Some lessons from Wikipedia. In *Proceedings - 2008 International MCETECH Conference on e-Technologies, MCETECH 2008*, pages 212–216, 2008.
- M. OLSON**, 1965. *The logic of Collective Action*. Harvard University Press, Cambridge Mass.
- S. O'MAHONY**, May 2007. The governance of open source initiatives: what does it mean to be community managed? *Journal of Management and Governance*, 11(2):139–150.

- S. O'MAHONY** and **B. A. BECHKY**, 2008. Boundary organizations: Enabling collaboration among unexpected allies. *Administrative Science Quarterly*, 53:422–459.
- S. O'MAHONY** and **F. FERRARO**, 2007. The emergence of governance in an open source community. *Academy of Management Journal*, 50:1059–1106.
- F. ORTEGA**, April 2009. Wikipedia: A quantitative analysis. Doctoral thesis, Universidad Rey Juan Carlos, Móstoles, Spain, 228 pages.
- F. ORTEGA** and **J. M. GONZALEZ BARAHONA**, 2007. Quantitative analysis of the wikipedia community of users. In *Proceedings of the 2007 international symposium on Wikis, WikiSym '07*, pages 75–86, New York, NY, USA. ACM.
- F. ORTEGA** and **D. IZQUIERDO-CORTAZAR**, 2009. Survival analysis in open development projects. In *Proceedings of the 2009 ICSE Workshop on Emerging Trends in Free/Libre/Open Source Software Research and Development, FLOSS '09*, pages 7–12, Washington, DC, USA. IEEE Computer Society.
- F. ORTEGA**, **D. IZQUIERDO-CORTAZAR**, **J. M. GONZALEZ-BARAHONA**, and **G. ROBLES**, January 2009. On the analysis of contributions from privileged users in virtual open communities. In *42th Hawaii International Conference on System Sciences (HICSS)*, page 110, Waikoloa, Big Island, Hawaii, USA. IEEE Computer Society.
- E. OSTROM**, 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press.
- K. PANCIERA**, **A. HALFAKER**, and **L. TERVEEN**, 2009. Wikipedians are born, not made: a study of power editors on Wikipedia. In *Proceedings of the ACM 2009 international conference on Supporting group work, GROUP '09*, pages 51–60, New York, NY, USA. ACM.
- T. K. PARK**, August 2011. The visibility of wikipedia in scholarly publications. *First Monday*, 16(8). URL: <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/3492/303>.
- G. PARMENTIER** and **T. PICQ**, 2006. Analyse exploratrice des pratiques de GRH innovantes du secteur du jeu vidéo. In *Proceedings. XVIIe Congrès de l'AGRH - Le travail au coeur de la GRH*, 2006.

- J. PÉNIN**, 2011. Open source innovation: Towards a generalization of the open source model beyond software. *Revue d'économie industrielle*, 136:65–88.
- C. PENTZOLD**, August 2011. Imagining the Wikipedia community: What do Wikipedia authors mean when they write about their "community"? *New Media & Society*, 13(5):704–721.
- G. PODERI**, 2009. Comparing featured article groups and revision patterns correlations in Wikipedia. *First Monday*, 14(5). URL: <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2365/2182>.
- M. E. PORTER** and **M. R. KRAMER**, December 2006. Strategy & Society. The Link Between Competitive Advantage and Corporate Social Responsibility. *Harvard Business Review*, pages 78–92.
- P. PRASARNPHANICH** and **C. WAGNER**, 2009. The role of wiki technology and altruism in collaborative knowledge creation. *Journal of Computer Information Systems*, 49(4):33–41.
- W. QU**, **Z. YANG**, and **Z. WANG**, 2011a. Multi-level framework of open source software adoption. *Journal of Business Research*, 64(9):997–1003.
- M. RAHMAN**, 2008. An Analysis of Wikipedia. *Journal of Information Technology Theory and Application*, 9:81.
- B. RAMDANI**, **P. KAWALEK**, and **O. LORENZO**, 2009. Predicting SMEs' adoption of enterprise systems. *Journal of Enterprise Information Management*, 22(1):10–24.
- S. RANSBOTHAM** and **G. C. KANE**, September 2011. Membership Turnover and Collaboration Success in Online Communities: Explaining Rises and Falls from Grace in Wikipedia. *MIS Quarterly*, 14(9): 541–545.
- L. H. RECTOR**, 2008. Comparison of Wikipedia and other encyclopedias for accuracy, breadth, and depth in historical articles. *Reference Services Review*, 36(1):7 – 22.
- T. S. ROBERTSON** and **H. GATIGNON**, July 1986. Competitive effects on technology diffusion. *Journal of Marketing*, 50(3):1–12.
- E. M. ROGERS**, 1983. *Diffusion of Innovations*. New York: The Free Press, 3rd edition.

- B. ROLANDSSON, M. BERGQUIST, and J. LJUNGBERG**, May 2011. Open source in the firm: Opening up professional practices of software development. *Research Policy*, 40(4):576–587. URL: <http://ideas.repec.org/a/eee/respol/v40y2011i4p576-587.html>.
- P. ROMER**, 1993. The economics of new ideas and new goods. In *Proceedings. Annual Conference on Development Economics, 1992*, Banque Mondiale, Banque Mondiale, Washington D. C., 1993.
- D. ROODMAN**, 2011. Fitting fully observed recursive mixed-process models with CMP. *The Stata Journal*, 11(2):159–206.
- M. ROSEMAN**, 1999. ERP software characteristics and consequences. In *Proceedings of the 7th European Conference on Information Systems*, Copenhagen.
- C. ROYAL and D. KAPILA**, 2009. What’s on Wikipedia, and What’s Not...? *Social Science Computer Review*, 27(1):138–148.
- F. RULLANI and S. HAEFLIGER**, 2013. The periphery on stage: The intra-organizational dynamics in online communities of creation. *Research Policy*, 42(4):941–953.
- D. E. RUPP, J. GANAPATHI, R. V. AGUILERA, and C. A. WILLIAMS**, June 2006. Employee Reactions to Corporate Social Responsibility: An organizational Justice Framework. *Journal of Organizational Behavior*, 27(4):537–543.
- A. SANTANA and D. WOOD**, 2009. Transparency and social responsibility issues for Wikipedia. *Ethics and Information Technology*, 11:133–144.
- W. SCACCHI**, 2007. Free/open source software development: recent research results and emerging opportunities. In *ESEC-FSE companion 07: The 6th Joint Meeting on European software engineering conference and the ACM SIGSOFT symposium on the foundations of software engineering*, pages 459–468, New York, NY, USA. ACM.
- A. SCHILLING, S. LAUMER, and T. WEITZEL**, January 2012. Who will remain? an evaluation of actual person-job and person-team fit to predict developer retention in floss projects. In *45th Hawai’i International Conference on System Sciences*, pages 3446–3455, January 2012.

- J. SCHROER** and **G. HERTEL**, 2009. Voluntary engagement in an open web-based encyclopedia: Wikipedians and why they do it. *Media Psychology*, 12(1):96–120.
- C. M. SCHWEIK**, 2006. Free/Open-Source Software as a Framework for Establishing Commons in Science. In (**HESS AND OSTROM, 2006C**), editor, *Understanding Knowledge as a Commons. From Theory to Practice*.
- P. B. SEDDON**, 1997. A respecification and extension of the DeLone and McLean model of IS success. *Information Systems Research*, 8(3):240–253.
- N. SERRANO** and **J. SARRIEL**, May-June 2006. Open source software ERPs: a new alternative for an old need. *Software, IEEE*, 23(3):94–97.
- S. K. SHAH**, July 2006. Motivation, Governance, and the Viability of Hybrid Forms in Open Source Software Development. *Management Science*, 52(2):1000–1014.
- W. G. SHEPHERD**, 1990. *The Economics of Industrial Organization*. Prentice Hall International Editions, London. 3rd edition.
- T. SIMCOE**, 2006. Open Standards and Intellectual Property Rights. In **CHESBROUGH ET AL. (2006)**, editor, *Open Innovation: Researching a New Paradigm*, pages 161–183.
- S. SLATTER**, 1992. *Gambling on Growth: How to Manage the Small High-Tech Firm*. Wiley, New York.
- R. SOEKATNO** and **A. GIRI**, 2005. The Indonesian language Wikipedia. In *Wikimania05*, 2005. URL: <http://meta.wikimedia.org/wiki/Transwiki:Wikimania05/Paper-RS1>.
- R. E. STAKE**, 2008. Case study. In (**DENZIN AND LINCOLN, 2008**), editor, *Strategies of qualitative inquiry*, pages 119–150.
- S. L. STAR** and **J. R. GRIESEMER**, August 1989. Institutional ecology, 'translations' and boundary objects: Amateurs and professionals in Berkeley's museum of vertebrate zoology, 1907-39. *Social Studies of Science*, 19(3):387–420.
- K. STEWART** and **S. GOSAIN**, 2006a. The impact of ideology on effectiveness in open source software development teams. *MIS Quarterly*, 30(2):291–314.

- J. K. STRATMAN**, 2007. Realizing benefits from enterprise resource planning: Does strategic focus matter? *Production and Operations Management*, 16(2):203–216.
- R. STROSS**, 2008. *Planet Google*. Free Press, New York.
- B. STVILIA, M. B. TWIDALE, L. C. SMITH, and L. GASSER**, 2008. Information Quality Work Organization in Wikipedia. *Journal of the American Society for Information Science and Technology*, 59(6):983–1001.
- E. SVOBODA**, May 2006. One-click content, no guarantees. *IEEE Spectrum*, 43:64–65. URL: <http://spectrum.ieee.org/telecom/internet/oneclick-content-no-guarantees>.
- E. B. SWANSON**, 1994. Information systems innovation among organizations. *Management Science*, 40(9):1069–1092.
- D. TEECE**, 1986. Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy*, 15 (7):285–305.
- D. TEECE, G. PISANO, and A. SHUEN**, 1997. Dynamic capabilities and strategic management. *Strategic Management Journal*, 18(7):509–533.
- L. THÉVENOT**, 2007. The plurality of cognitive formats and engagements: moving between the familiar and the public. *European Journal of Social Theory*, 10(3):413–427.
- J. TIROLE**, 1989b. *The Theory of Industrial Organization*. MIT Press.
- L. G. TORNATSKY and K. J. KLEIN**, February 1982. Innovation characteristics and innovation adoption implementation. *IEEE Transactions on Engineering Management*, 29(1):28–45.
- B. UZZI**, June 2008. A social network’s changing statistical properties and the quality of human innovation. *Journal of Physics A: Mathematical and Theoretical*, 41(22):224023, 12pgs.
- B. UZZI and J. SPIRO**, September 2005. Collaboration and creativity: The small world problem. *American Journal of Sociology*, 111(2):447–504.
- Z. VAN DIJK**, 2009. Wikipedia and lesser-resourced languages. *Language Problems and Language Planning*, 33(3):234–250.

- H. R. VARIAN**, 2005. *Intermediate microeconomics: a modern approach*. W. W. Norton, New York, NY, 2nd edition.
- V. VENKATESH, M. G. MORRIS, G. B. DAVIS**, and **F. DAVIS**, 2003. User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3):425–478.
- M. VICENTE**, 2008. The impact of companies on the dynamic of open source software development networks: Convergence or differentiation of production models and labor division? In *doctoriales du GDR TIC et société*, Rennes.
- F. B. VIÉGAS, M. WATTENBERG, J. KRISS**, and **F. VAN HAM**, 2007a. Talk before you type: Coordination in Wikipedia. In *Proceedings of HICSS 2007*, pages 78–87, 2007a.
- F. B. VIÉGAS, M. WATTENBERG**, and **M. M. MCKEON**, 2007b. The hidden order of Wikipedia. In *Proceedings of the 2nd international conference on Online communities and social computing, OCSC'07*, pages 445–454, Berlin, Heidelberg. Springer-Verlag.
- E. VON HIPPEL**, July 1986. Lead users: a source of novel product concepts. *Management Science*, 32(7):791–805.
- E. VON HIPPEL**, 1988. *The Sources of Innovation*. Oxford University Press, New York.
- E. VON HIPPEL**, 2002. Open source software as horizontal innovation networks- by and for users. *MIT Sloan School of Management W.P*, 4366-02.
- E. VON HIPPEL** and **G. VON KROGH**, March-April 2003. Open source software and the "private-collective" innovation model: Issues for organization science. *Organization Science*, 14(2):209–223.
- G. VON KROGH, S. HAEFLIGER, S. SPAETH**, and **M. W. WALLIN**, 2012. Carrots and Rainbows: Motivation and Social Practice in Open Source Software Development. *MIS Quarterly*, 36(2):649–676.
- G. VON KROGH, S. SPAETH**, and **K. R. LAKHANI**, 2003. Community, joining, and specialization in open source software innovation: A case study. *Research Policy*, 32(7):1217–1241.
- J. VOSS**, 2005. Measuring Wikipedia. In *Proceedings of the International Conference of the International Society for Scientometrics and Informetrics (ISSI)*, 10, Stockholm.

- D. WALDMAN, M. SULLY DE LUQUE, N. WASHBURN, and R. HOUSE**, 2006. Cultural and leadership predictors of corporate social responsibility values of top management: A GLOBE study of 15 countries. *Journal of International Business Studies*, 37:823–837.
- M. M. WASKO and S. FARAJ**, 2005. Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice. *MIS Quarterly*, 29(1):35–57.
- N. L. WATERS**, September 2007. Why you can't cite Wikipedia in my class. *Communications of the ACM*, 50:15–17.
- H. T. WELSER, D. COSLEY, G. KOSSINETS, A. LIN, F. DOKSHIN, G. GAY, and M. SMITH**, 2011. Finding social roles in Wikipedia. In *Proceedings of the 2011 iConference*, iConference '11, pages 122–129, New York, NY, USA. ACM.
- E. WENGER**, 1998. *Communities of practice: Learning, Meaning, and Identity*. Cambridge University Press, Cambridge, Mass.
- J. WEST**, 2003. How open is open enough? Melding proprietary and open source platform strategies. *Research Policy*, 32 (7):1259–1285.
- K. WEST and J. WILLIAMSON**, 2009. Wikipedia: Friend or Foe? *Reference Services Review*, 37(3): 260 – 271.
- D. M. WILKINSON and B. A. HUBERMAN**, April 2007. Assessing the value of cooperation in Wikipedia. *First Monday*, 12(4). URL: <http://www.firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1763/>.
- H.-L. YANG and C.-Y. LAI**, 2010. Motivations of Wikipedia content contributors. *Computers in Human Behavior*, 26(6):1377 – 1383.
- R. K. YIN**, 2009. *Case study research: design and methods*. SAGE, 4th edition.
- X. ZHANG and F. ZHU**, 2011. Group size and incentives to contribute: A natural experiment at chinese wikipedia. *The American Economic Review*, 101(4):1601–1615.

- X. ZHAO** and **M. J. BISHOP**, October 2011. Understanding and supporting online communities of practice: lessons learned from wikipedia. *EDUCATIONAL TECHNOLOGY RESEARCH AND DEVELOPMENT*, 59(5):711–735.
- J.-B. ZIMMERMANN**, September 1995a. Le concept de grappes technologiques. Un cadre formel. *Revue économique*, 46(5):1263–1295.
- V. ZLATIĆ** and **H. STEFANCIĆ**, 2011. Model of Wikipedia growth based on information exchange via reciprocal arcs. *EPL*, 93(5):58005.

6.2 References Written by Nicolas Jullien.

6.2.1 Cited in the text.

- K. CROWSTON**, **F. DELTOUR**, and **N. JULLIEN**, May 2013a. Open Source Software Adoption: A Technological Innovation Perspective. In *colloque annuel de l'Association Information et Management (AIM)*, May 2013a. URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2244222.
- K. CROWSTON**, **N. JULLIEN**, and **F. ORTEGA**, January 2013b. Is Wikipedia Inefficient? Modelling Effort and Participation in Wikipedia. In *Forty-sixth Hawai'i International Conference on System Sciences (HICSS-46)*. IEEE Computer society, January 2013b. URL: <http://ssrn.com/abstract=1960696>.
- J.-M. DALLE** and **N. JULLIEN**, January 2003. 'Libre' Software: Turning Fads into Institutions? *Research Policy*, 32(1):1–11.
- S. DEJEAN** and **N. JULLIEN**, 2012. Enrolled since the beginning. Assessing Wikipedia Contributors' Behavior by Their First Contribution. In *Working paper*. M@rsouin, 2012.
- D. DEMAZIÈRE**, **F. HORN**, and **N. JULLIEN**, June 2005. Le travail des développeurs de logiciels libres. La mobilisation dans des "communautés distantes". *Cahiers lillois d'économie et de sociologie*, 46: 171–194.
- N. JULLIEN**, 2003. Le marché francophone du logiciel libre. *Systèmes d'information et management*, 8 (1):77–100.

- N. JULLIEN**, 2012. What We Know About Wikipedia: A Review of the Literature Analyzing the Project(s). In *Working paper*. M@rsouin, Télécom Bretagne, 2012.
- N. JULLIEN, M. CLÉMENT-FONTAINE, and J.-M. DALLE**, 2002a. New economic models, new software industry economy. Technical report, RNTL (French National Network for Software Technologies) project, 202 pages. URL: http://www.marsouin.org/IMG/pdf/fichier_rapporte-3.pdf.
- N. JULLIEN, M. CLÉMENT-FONTAINE, and J.-M. DALLE**, 2002b. Nouveaux modèles économiques, nouvelle économie du logiciel. Technical report, Projet RNTL, 218 pages. URL: http://www.marsouin.org/article.php3?id_article=78.
- N. JULLIEN and J. PÉNIN**, 2013. Innovation ouverte: vers la génération 2.0. In **J.-P. DENIS, T. HAFSI, A. C. MARTINET, and F. TANNERY**, editors, *Encyclopédie de la Stratégie*. Economica. to be published.
- N. JULLIEN and K. ROUDAUT**, 2012. Can open source projects succeed when the producers are not users? lessons from the data processing field. *Management international = International management = Gestión internacional*, 16:113–127.
- N. JULLIEN, K. ROUDAUT, and S. LE SQUIN**, novembre-décembre 2011. L’engagement dans des collectifs de production de connaissance en ligne. Le cas GeoRezo. *Revue française de socio-économie*, 8(2):59–83.
- N. JULLIEN and J.-B. ZIMMERMANN**, 2002. Le logiciel libre: une nouvelle approche de la propriété intellectuelle? *Revue d’économie industrielle*, 99:159–178. Special Issue, *Les droits de propriété intellectuelle: nouveaux domaines, nouveaux enjeux*.
- N. JULLIEN and J.-B. ZIMMERMANN**, 2006b. New approaches to intellectual property: from open software to knowledge based industrial activities. In **S. LABORY and P. BIANCHI**, editors, *Industrial Handbook on Industrial Policy*, pages 243–264. Edward Elgar (EE).
- N. JULLIEN and J.-B. ZIMMERMANN**, 2006c. Peut-on envisager une écologie du logiciel libre favorable aux nuls? *Terminal*, 97-98:33–47. URL: http://www.marsouin.org/article.php3?id_article=65.

- N. JULLIEN** and **J.-B. ZIMMERMANN**, 2009. Firms' contribution to open-source software and the dominant user's skill. *European Management Revue*, 6:130 – 139.
- N. JULLIEN** and **J.-B. ZIMMERMANN**, 2011a. Floss firms, users and communities: a viable match? *Journal of Innovation Economics*, 1(7):31–53.
- N. JULLIEN** and **J.-B. ZIMMERMANN**, 2011b. Floss in an industrial economics perspective. *Revue d'économie industrielle*, 136:39–64.

6.2.2 Documents written by Nicolas Jullien not cited in the text.

- M. CLÉMENT-FONTAINE** and **N. JULLIEN**, 2002. Le contexte économique et juridique de l'émergence d'une nouvelle économie du logiciel. In (**JULLIEN ET AL., 2002B**), editor, *Nouveaux modèles économiques, nouvelle économie du logiciel*, pages 7–29.
- J.-M. DALLE** and **N. JULLIEN**, 2000. Nt vs. linux, or some explorations into the economics of free software. In **G. BALLOT** and **G. WEISBUCH**, editors, *Applications of Simulation to Social Sciences*, pages 399–416. Hermès, Paris.
- D. DEBOIS**, **N. JULLIEN**, **T. PÉNARD**, **A. POULAIN-MAUBANT**, **J. VÉTOIS**, and **J.-B. ZIMMERMANN**, Fall-Winter 1999. Logiciels libres: de l'utopie au marché. *Terminal, special issue*, 80-81.
- D. DEMAZIÈRE**, **F. HORN**, and **N. JULLIEN**, 2006. How free software developers work: The mobilization of 'distant communities'. *SSRN eLibrary*. URL: <http://ssrn.com/abstract=1301572>.
- J.-P. DURAND** and **N. JULLIEN**, 2012. Regards sur le travail et les tic, entrevue avec jean-pierre durand. *Terminal*, 110.
- N. JULLIEN**, 1999. Linux: la convergence du monde Unix et du monde PC. *Terminal*, 80/81:43–70. Special Issue, *Le logiciel libre*.
- N. JULLIEN**, November 2001. *Impact du logiciel libre sur l'industrie informatique*. PhD, Université de Bretagne Occidentale / ENST Bretagne, mention: sciences économiques, 307 pages. URL: http://www.marsouin.org/article.php3?id_article=82.

- N. JULLIEN**, June 2005. Firms' strategies facing floss diffusion. *Upgrade, The European Journal for the Informatics Professional*, VI-3. URL: <http://www.upgrade-cepis.org/issues/2005/3/upgrade-vol-VI-3-low-res.pdf>.
- N. JULLIEN**, 2008. Developing 'Free/Libre/Open Source Software (FLOSS)' - A Market Driven Investment. *SSRN eLibrary*. URL: <http://ssrn.com/paper=1298954>.
- N. JULLIEN**, 2009a. A historical analysis of the emergence of free cooperative software production. In **P. F. TIAKO**, editor, *Software Applications: Concepts, Methodologies, Tools, and Applications (6 Volumes)*. IGI Global, Hershey, PA, USA.
- N. JULLIEN**, 2009b. A historical analysis of the emergence of free cooperative software production. In **M. PAGANI**, editor, *Encyclopedia of multimedia technology and networking, second edition*, pages 605–612. IGI Global, Hershey, PA, USA.
- N. JULLIEN** and **M. CLÉMENT-FONTAINE**, 2002a. The economic and legal context of a new software economy. In (**JULLIEN ET AL., 2002A**), editor, *New Economic Models, New Software Industry Economy*, pages 1–24.
- N. JULLIEN** and **M. CLÉMENT-FONTAINE**, 2002b. Le contexte économique et juridique de l'émergence d'une nouvelle économie du logiciel. In (**JULLIEN ET AL., 2002A**), editor, *Nouveaux modèles économiques, nouvelle économie du logiciel*, pages 7–30.
- N. JULLIEN**, **J. MUÑOZ**, and **M. BURNIER**, 2012. Dossier travail et tic. *Terminal*, 110.
- N. JULLIEN** and **D. PHAN**, November 1999. From the 'unix world' towards the 'linux community': An historical co-evolutionary perspective on 'small worlds'. In *EAEPE proceedings*, November 1999.
- N. JULLIEN** and **J.-B. ZIMMERMANN**, July 2006a. Les réussites économiques du libre. *Revue Télécom. Numéro spécial: les logiciels libres*, 146:19–23.
- N. JULLIEN** and **J.-B. ZIMMERMANN**, 2013. Le logiciel libre, un renouveau du modèle industriel coopératif de l'informatique. In **C. PALOQUE-BERGES** and **C. MASUTTI**, editors, *Histoires et cultures du Libre*, pages 135–164. Framabook.

- K. ROUDAUT** and **N. JULLIEN**, 2013. Des usages professionnels individualisés des outils de réseau social, révélateurs d'une désaffiliation identitaire de l'entreprise? *Revue Française de socio-économie*. Soumis.

Abstract

This document presents the research I have undertaken over the last decade. It is both retrospective and prospective in the sense that, although it is obviously focused on my past activities, it also indicates ways for future research. The main topic of my overall research can be summarized as follows: I explore the development of online, open projects, or communities of creation, such as Free, Libre, Open Source Software (FLOSS), from an economics point of view. This means that in addition to renewing the answers to Olson's question about the individual participation to collective action (1965), it questions also the why and how companies participate in this process, renewing Arrow's dilemma (1962) on the incentives to produce innovation and the incentive to disseminate this innovation, and the way people organize themselves to transform participation into concrete pieces of knowledge, being software or encyclopedia articles.